AEDC-TR-94-17

Spectral Analysis Techniques Using Wavelets As An Alternative to Fourier Analysis For Transient Dynamic Data

Kenneth R. Kimble
The University of Tennessee Space Institute

Thomas F. Tibbals
Sverdrup Technology, Inc., AEDC Group

January 1995

Final Report for Period October 1993 — September 1994

ARNOLD ENGINEERING DEVELOPMENT CENTER
ARNOLD AIR FORCE BASE, TENNESSEE
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICES

When U. S. Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise, or in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Qualified users may obtain copies of this report from the Defense Technical Information Center.

References to named commercial products in this report are not to be considered in any sense as an endorsement of the product by the United States Air Force or the Government.

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.
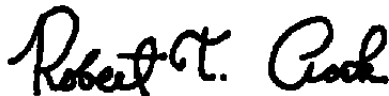
APPROVAL STATEMENT

This report has been reviewed and approved.

JAMES D. MITCHELL
Propulsion Technology
Applied Technology Division
Test Operations Directorate

Approved for publication:

FOR THE COMMANDER

ROBERT T. CROOK
Assistant Chief, Applied Technology Division
Test Operations Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE January 1995	3. REPORT TYPE AND DATES COVERED Final Report for Period October 1993 - September 1994

4. TITLE AND SUBTITLE

Spectral Analysis Techniques Using Wavelets as an Alternative to Fourier Analysis for Transient Dynamic Data

5. FUNDING NUMBERS

Program Element 65807F

6. AUTHOR(S)

Kimble, Kenneth R., University of Tennessee Space Institute, and Tibbals, Thomas F. Sverdrup Technology, Inc., AEDC Group

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Arnold Engineering Development Center/DOT
Air Force Materiel Command
Arnold Air Force Base, TN 37389-9011

8. PERFORMING ORGANIZATION (REPORT NUMBER)

AEDC-TR-94-17

9. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)

Arnold Engineering Development Center/DO
Air Force Materiel Command
Arnold Air Force Base, TN 37389-9010

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

Available in Defense Technical Information Center (DTIC).

12A. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12B. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Various dynamic data analysis tools for one-dimensional time-domain signals are employed to determine the frequency content of a signal for mechanical analysis. When tied to the fundamental frequency of the various components comprising the machinery being evaluated, this information gives an indication of the state or health of the machine. Current techniques for evaluating dynamic data for potential mechanical problems are primarily centered around the Fast Fourier Transform (FFT) and the Short Time Fourier Transform (STFT). However, the use of Fourier analysis for frequency component extraction is restricted to bandlimited stationary signals. Because of the stationary requirement, small transients may not be detected due to a smoothing effect of the FFT, or the FFT spectrum may be smeared due to frequency ramping and abrupt incidents or discontinuities in the signal. Various techniques have been employed to overcome the limitations of the FFT for non-stationary data. These techniques include windowed Fourier analysis (including the STFT) and a background in wavelet theory based upon the analyzing function basis approach so that wavelet theory may be contrasted against Fourier analysis. A model signal with stationary and transient characteristics is developed to permit comparisons of various analysis techniques based upon a known analytic signal which resembles a real vibration signal. Some applications of wavelets to other transient signals are also provided.

14. SUBJECT TERMS

Fourier, FFT, wavelets, Daubechies, Malvar

15. NUMBER OF PAGES

67

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT SAME AS REPORT

PREFACE

This report provides an overview of some advanced digital signal processing techniques for one- and two-dimensional data that may be useful for pattern recognition, real-time image processing, and on-line test article health monitoring functions. Topics include a review of classical Fourier analysis and its application to transient data, the Short Time Fourier Transform, multi-resolution analysis, wavelet analysis, and applications of wavelet analysis to some typical AEDC-type signals.

CONTENTS

ILLUSTRATIONS

1.0 INTRODUCTION

Various dynamic data analysis tools for one-dimensional time-domain signals are employed to determine the frequency content of a signal for mechanical analysis. When it is tied to the fundamental frequency of the various components comprising the machinery being evaluated, this information gives an indication of the state or health of the machine. Current techniques for evaluating dynamic data for potential mechanical problems are primarily centered around the Fast Fourier Transform (FFT) and the Short Time Fourier Transform (STFT). However, the use of Fourier analysis for frequency component extraction is restricted to bandlimited stationary signals. Because of this restriction, small transients may not be detected due to a smoothing effect of the FFT, or the FFT spectrum may be smeared due to frequency ramping and abrupt incidents or discontinuities in the signal. Various techniques have been employed to overcome the limitations of the FFT for non-stationary data. These techniques include windowed Fourier Transform (Gabor or Short Time Fourier Transform), synchronous sampling to remove RPM-ramp effects, Wigner-Ville analysis, and, more recently, wavelet analysis. This report provides an overview of Fourier analysis (including the STFT) and a background in wavelet theory based upon the analyzing function basis approach so that wavelet theory may be contrasted against Fourier analysis. A model signal with stationary and transient characteristics is developed to permit comparisons of various analysis techniques based upon a known analytic signal which resembles a real vibration signal. Some applications of wavelets to other transient signals are also provided.

2.0 FOURIER TRANSFORM

Consider the representation of a finite-power signal, $x(t)$, defined on the interval $(t_0, t_0 + T)$ in terms of a set of preselected time functions $\Phi_1(t)$, ..., $\Phi_n(t)$. It is convenient to choose these functions with properties analogous to the orthogonal unit vectors of Cartesian space, i.e., orthogonal such that

$$\int_{t_0}^{t_0+T} \Phi_m(t)\Phi_n^*(t)dt = 0 \quad \forall m \neq n$$

and normalized such that

$$\frac{1}{T}\int_{t_0}^{t_0+T} |\Phi_n(t)\Phi_n^*(t)|dt = 1$$

where $\Phi_n^*(t)$ is the complex conjugate of $\Phi_n(t)$. Therefore, these functions are orthonormal functions on the interval $(t_0, t_0 + T)$, where the orthonormality condition was chosen since the signal, being periodic, is bounded in power though unbounded in energy.

Now consider an approximation $\tilde{x}(t)$ to $x(t)$ based upon a series expansion of the form

$$\tilde{x}(t) = \sum_{n=1}^{N} C_n \Phi_n(t)$$

where N is a given finite positive integer and where the constants C_n are to be chosen such that $\tilde{x}(t)$ represents $x(t)$ as closely as possible. A criterion of measure for this *closeness* of approximation must be chosen based upon the signal's characteristics. In this case, we choose to measure the error in a square-integral sense since our concern is with power signals. We define this error as

$$\epsilon_N = \int_{t_0}^{t_0+T} |x(t) - \tilde{x}(t)|^2 dt ,$$

and we wish to find the C_n coefficients to minimize this error. Substituting for $\tilde{x}(t)$ yields

$$\epsilon_N = \int_{t_0}^{t_0+T} \left[x(t) - \sum_{n=1}^{N} C_n \Phi_n(t) \right] \left[x^*(t) - \sum_{n=1}^{N} C_n^* \Phi_n^*(t) \right] dt$$

$$\epsilon_N = \int_{t_0}^{t_0+T} |x(t)|^2 dt - \sum_{n=1}^{N} \left[C_n^* \int_{t_0}^{t_0+T} x(t)\Phi_n^*(t)dt + C_n \int_{t_0}^{t_0+T} x^*(t)\Phi_n(t)dt \right] dt + \sum_{n=1}^{N} |C_n|^2 \} .$$

Adding and subtracting $\sum_{n=1}^{N} \left| C_n^* \int_{t_0}^{t_0+T} x(t)\Phi_n^*(t) \right|^2 dt$ yields

$$\epsilon_N = \int_{t_0}^{t_0+T} |x(t)|^2 dt - \sum_{n=1}^{N} \left| \int_{t_0}^{t_0+T} x(t)\Phi_n^*(t) \right|^2 dt + \sum_{n=1}^{N} \left| C_n - \int_{t_0}^{t_0+T} x(t)\Phi_n^*(t)dt \right|^2 dt .$$

Now since the first two terms above are independent of the C_n coefficients of our approximation, and the last term is non-negative and adds to these two terms, then the only way to minimize the error is to choose the approximation coefficients C_n such that

$$C_n = \int_{t_0}^{t_0+T} x(t)\Phi_n^*(t)dt \quad n = 1, 2, ..., N$$

Note that this is an inner product such that $C_n = <x, \Phi_n>$. For special choices of the set of functions $\Phi_1(t), ..., \Phi_N(t)$, referred to as complete sets in the Hilbert function space L^2 of all square-integrable functions, it will be true that

$$\lim_{N \to \infty} \epsilon_{N,min} = \lim_{N \to \infty} \left[\int_{t_0}^{t_0+T} |x(t)|^2 dt - \sum_{n=-M}^{N} |C_n|^2 \right] = 0$$

6

for any signal that is square-integrable, i.e.,

$$\int_{t_0}^{t_0+T} |x(t)|^2 dt < \infty .$$

In the case of zero square-integral error,

$$x(t) = \sum_{n=1}^{\infty} C_n \Phi_n(t)$$

and

$$\frac{1}{T} \int_{t_0}^{t_0+T} |x(t)|^2 dt = \sum_{n=1}^{\infty} |C_n|^2 .$$

This is the generalized Parseval theorem which states that the mean power of the periodic signal is the sum of the squares of the approximating coefficients. The exponential functions of the Fourier series form a complete orthogonal set in L^2 such that

$$\Phi_n(t) = e^{inw_o t} , \qquad n = 0, \pm 1, \pm 2, ..., \infty$$

are axes of L^2 and Fourier analysis projects $x(t)$ onto these axes just as in the case of breaking a vector into its orthogonal components is to project the vector onto the orthogonal axes of Cartesian space. Now $w_o = 2\pi/T$ where T is the *fundamental* period, that is, T is the minimum period for a complete sine wave of the lowest frequency of interest. Thus, partial sums of exponential Fourier series minimize the square-integrable error between the series and the signal under investigation (Ref. 1).

Since sine waves have a period of 2π, a function $x(t)$ that is square integrable (i.e., piecewise continuous) on the interval $(0, 2\pi)$ i.e.,

$$\frac{1}{2\pi} \int_{0}^{2\pi} |x(t)|^2 dt < \infty ,$$

is an element of the Hilbert function space $L^2(0, 2\pi)$ of all square integrable functions defined on $(0, 2\pi)$ and expansions and contractions of $x(t)$ (i.e., varying the number of cycles on the interval $(0, 2\pi)$) form a basis of the $L^2(0, 2\pi)$ function space. By *basis* is meant a set of linearly independent functions which can be combined to form another function which also exists in the *same* function space. Then $x(t)$ has a Fourier series representation:

$$x(t) = \sum_{n=-\infty}^{\infty} C_n e^{inw_o t}$$

7

where, for $n = -\infty \ldots, -2, -1, 0, 1, 2, \ldots \infty$,

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} x(t) e^{-in\omega_0 t} dt .$$

Notice that the number of basis functions is infinite for this representation.

Per Chui (Ref. 2), this series converges in L^2 such that

$$\lim_{M,N \to \infty} \frac{1}{2\pi} \int_0^{2\pi} |x(t) - \sum_{n=-M}^{N} C_n e^{in\omega_0 t}|^2 dt = 0 .$$

There are three distinct features of $x(t)$ in the Fourier series representation. First, $x(t)$ is decomposed into mutually orthogonal components $C_n e^{i\omega_0 t}$ since the inner product

$$\frac{1}{2\pi} \int_0^{2\pi} C_n e^{in\omega_0 t} C_m e^{-im\omega_0 t} dt = \begin{cases} 0 & m \neq n \\ 1 & m = n \end{cases} .$$

Second, the bases of this orthonormal representation are all generated from a single function $\omega(t) = e^{i\omega_0 t}$ by integral dilations, i.e., $w_n(t) = w(nt)$ for all integers n. Finally, the Fourier expansion coefficients are determined via an inner product $c_n = <x, w_n>$. "It is orthogonality that allows us to find each term separately [no cross-terms means no cross-talk or interdependence between frequencies], and it is completeness that allows the sines and cosines to reproduce $x(t)$." (Ref. 3).

However, the use of this complex exponential basis is not without complications. First, the approximation using Fourier coefficients is an infinite summation which can only be made finite by bandlimiting the signal. Second, the basis functions themselves have infinite extent in the time domain, and therefore cannot readily approximate a short-lived event in the signal. Third, the signal must be periodic, or at least the results of Fourier analysis assume periodicity, which may not be the case for typical *real* signals. Finally, the signal should meet the Dirichlet conditions to ensure the existence of the Fourier series approximation[1] (i.e., the series converges to the continuous signal $x(t)$). The Dirichlet conditions are:

---

1. Note that if these conditions hold as $T \to \infty$, i. e., the signal is *aperiodic*, then the signal's Fourier Integral Transform exists and the signal can be approximated by a sum of pure sinusoids which vary *continuously* in frequency. In this case the spectrum is continuous instead of discrete as in the Fourier Series representation.

- If the function has discontinuities, their number must be finite in any period.

- The function must contain a finite number of maxima and minima during any period.

- The function must be absolutely integrable in any period, i.e.,

$$\int_0^T |x(t)|dt < \infty$$

where $x(t)$ is a continuous ($\forall t$, at least piecewise) describing function which approximates the actual (real) signal. Note that the Dirichlet conditions impose *stationarity* onto the signal. In addition, note that there is no time domain translation parameter, as will be seen later for wavelets, since it makes no sense to translate an infinite extent function — this is handled by the concept of phase. However, this implies no time domain localization is possible, and consequently, the Fourier coefficients are independent of time.

For a sampled (digital) signal $x_s(n\Delta t)$ as an approximation to $x(t)$ (where Δt = sample interval and $T = N\Delta t$ analysis time block), we can then approximate c_k as

$$C_k \cong \frac{1}{T}\sum_{n=0}^{N-1}\left(x_s(n\Delta t)e^{-ik\omega_0 n\Delta t}\right)\frac{T}{N}$$

$$= \frac{1}{N}\sum_{n=0}^{N-1}x_s(n\Delta t)e^{-ik\omega_0 \frac{n}{N}T}$$

$$= \frac{1}{N}\sum_{n=0}^{N-1}x_n e^{-ik\omega_0 \frac{n}{N}T}$$

where $t \to n\Delta t$ and $\Delta t \to T/N$, and N is the number of samples (measurement values) of the signal within a given localized time interval of the sampling window. This defines the Discrete Fourier Transform (DFT) for a bandlimited signal. The Fast Fourier Transform (FFT) is merely a fast algorithm for performing the DFT.

3.0 SHORT TIME FOURIER TRANSFORM (GABOR TRANSFORM)

To better approximate transient, short-lived or time-localized phenomena, we must *fix* the Fourier analysis beyond that of just the simple windows caused by digitization of the signal or the accumulation of N samples for use in the FFT. One approach is to window the signal in time to better approximate the local characteristics. The Short Time Fourier Transform (STFT, or *Gabor Transform*) attempts to provide this localization. The STFT of signal $x(t)$ is defined as

9

$$STFT(t,\omega) = \int_{-\infty}^{\infty} x(\zeta - t)g(\zeta)e^{-in\omega\zeta}d\zeta$$

where the Gaussian window $g(\zeta)$, $\{\zeta : [0, \tau]\}$, centered about $\tau/2$ is defined as

$$g(\zeta) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(\zeta - \frac{\tau}{2})^2}{2\sigma^2}} ,$$

and σ (standard deviation) is the inflection point of concavity of the curve such that large σ indicates a flat curve and small σ indicates a peaked curve. For comparison purposes with wavelets later, let us define the window $g(\zeta)$ to include the transformation exponential such that

$$g(\zeta) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(\zeta - \frac{\tau}{2})^2}{2\sigma^2}}e^{-in\omega\zeta}$$

essentially windows the transform basis functions. Then the STFT is again an inner product, $<x, g_{\omega,\tau}>$, and linearity is preserved. Per Chui (Ref. 2), the time window width is $\Delta g_\sigma(t) = \sqrt{2}\sigma$ while the frequency (spectral) window width is $\Delta g_\sigma(\omega) = \sqrt{2}/\sigma$, and these window widths are *constant* for the analysis independent of frequency, as shown in the STFT calculation on the next page. Note that the product of these window widths is a constant in agreement with the Heisenberg Uncertainty Principle.
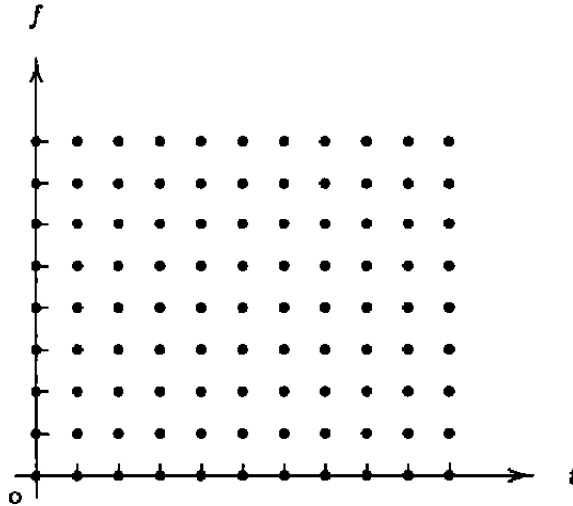
STFT CALCULATION GRID

The STFT can be discretized following the same approach as for the Discrete Fourier Transform. For a sampled (digital) signal $x_s(n\Delta t)$ as an approximation to $x(t)$ (where Δt = sample interval and $T = N\Delta t$ = analysis time block), we can then approximate $c_{k,m}$ for $\zeta = m\Delta t$, $m < N$, as

$$
\begin{aligned}
c_{k,m} &\cong \sum_{n=0}^{N-1} \left(x_s(m\Delta t - n\Delta t) g(m\Delta t) e^{-ik\omega_0 m\Delta t} \right) \frac{T}{N} \\
&= \frac{T}{N} \sum_{n=0}^{N-1} x_s(m\Delta t - n\Delta t) g(m\Delta t) e^{-ik\omega_0 \frac{n}{N} T} \\
&= \frac{T}{N} \sum_{n=0}^{N-1} x(m - n) g_m e^{-ik\omega_0 \frac{n}{N} T}
\end{aligned}
$$

where again t $\rightarrow n\Delta t$ and $\Delta t \rightarrow T/N$, N is the number of samples (measurement values) of the signal within a given localized time interval of the sampling window, and M is the number of samples within the Gaussian window $g_m = 1/\sigma \sqrt{2\pi} e^{-(m-M/2)^2/2\sigma^2} e^{-in\omega m}$. Note that the STFT actually has two windows applied in the time domain. The first window is the blocking of the data itself, just as in the normal DFT. The second window is the Gaussian window which localizes the signal in time by making the signal look as if it has compact support. However, one must remember that this *compact support* is assumed to be periodic. In addition, the window must be sufficiently narrow such that the signal is **stationary** within the window so that the Fourier Transform is applicable. Obviously, the STFT will work for a narrowband signal such that an appropriate window can be satisfactorily narrowed without adversely affecting the lower frequencies.

3.1 APPLICATION TO MODEL SIGNAL

The use of the STFT for gear fault detection via vibrational analysis is described in Ref. 4. Gear fault detection is an inherently transient analysis since a problem with a gear tooth can only be perceived during the time the tooth is in mesh with other components. The STFT time-frequency output for such an analysis, however, requires some interpretation. Following the example of Ref. 4, consider a model signal which resembles a gear fault and is composed of two sinusoids with stationary properties, another which is amplitude modulated, and two Gaussian-shaped impulses as described by

$$
x(n) = A_1 \sin \frac{2\pi f_1 n}{N} + \left(1 + A_2 \cos \frac{2\pi f_2 n}{N} \right) A_3 \sin \frac{2\pi f_0 n}{N} + A_4 e^{\frac{B_1^2(n-n_1)^2}{N}} + A_5 e^{\frac{B_2^2(n-n_2)^2}{N}}
$$

for $0 \leq n < N - 1$.

The parameters used were: $N = 1024$, $f_0 = 10$ Hz, $f_1 = 80$ Hz, $f_2 = 128$ Hz, $A_1 = 1.5$, $A_2 = 0.75$, $A_3 = 1.0$, $A_4 = 3.0$, $A_5 = 3.0$, $B_1 = 64$, $B_2 = 64$, $n_1 = 320$, and $n_2 = 576$ and the sample rate at $2.56F_{max} = 327.68$ Hz (where F_{max} is the maximum static frequency component). Gaussian impulses were used to simulate a real mechanical transient event rather than delta functions. The Gaussian impulses have more than one sample describing the event as opposed to the single sample of a delta function. This suggests a more real energy distribution in time for a mechanical system, since few mechanical systems can instantaneously respond and decay as implied by a delta function. Figure 1 shows the signal and its spectrum. Note that there is no indication of the Gaussian-shaped impulses, and the amplitude modulated signal component at 128 Hz has split into two components centered around 128 Hz separated by twice the modulation frequency of $f_0 = 10$ Hz.

Applying the STFT to this signal with a window width of 32 points and no overlap results in the spectrogram of Fig. 2. The stationary sinusoids appear as straight horizontal bands in the time-frequency plot, continuous for the whole time interval. Note that the frequency resolution is coarse due to poor frequency localization at the low number of FFT points, while the time resolution adequately shows the amplitude modulation at $f = 128$, but scalloping has smeared the true amplitudes. The Gaussian impulses have lost their time dependence, and only contribute to an amplitude modulation effect for the 10-Hz component, since the impulses span more than one FFT window.

Now if the window width is changed to 128 points to provide better frequency localization, then the spectrogram of Fig. 3 results. Note that the time resolution has worsened in an attempt to better localize the frequency components. In this case, the Gaussian impulses are visible since the 10-Hz peak is sharper due to narrower frequency bin spacing. The low-frequency smearing for time slices three and five are the Gaussian impulses at $n = 320$ and $n = 576$, respectively. However, note that the amplitude modulation at 128 Hz is no longer discernible, and that it is impossible to determine whether the spikes at 118 Hz and 138 Hz are real components or a 10-Hz amplitude modulation of a 128-Hz component. To improve time resolution with the same size FFTs, the signal was analyzed with overlapping windows. Figures 4 and 5 illustrate the signal and its spectrum for 50-percent overlap with 32- and 128-point FFTs, respectively. Note that there is no change in the 50-percent overlap spectrogram for the 128-point FFTs as compared to the zero overlap spectrogram (Fig. 3). This is due to the signal characteristics being of smaller scale than what even the 50-percent overlapping provides. However, the 32-point FFT spectrogram with 50-percent overlap was able to discern the Gaussian peaks which were not discernible with zero overlap (Fig. 2). This improvement in information content via overlapping windows is at the heart of the Malvar wavelet approach, and will be discussed further in Section 5.0.

Again, a *constant* window width was utilized for all frequencies processed by the STFT, regardless of overlapping. As we shall see with wavelets, the ability to vary the window width with frequency will improve the analysis (Ref. 2). Also note that by changing window widths in a series of

such analyses, we can enhance various signal characteristics with subsequent trade-offs in time-frequency localizations for various components. This series of analyses constitutes a Multi-Resolution Analysis (MRA) as discussed in the next section. A multi-resolution analysis will allow us to adjust our windows in both the time and frequency domains to optimize time-frequency localization.

4.0 DAUBECHIES WAVELET TRANSFORM

4.1 THE MULTI-RESOLUTION CONCEPT

Multi-Resolution Analysis is a processing technique that adapts to the frequency range of interest to optimize the resolution in both the time and frequency domains. Generally, while maintaining the constraints of the Heisenberg Uncertainty Principle, this analysis adapts to the frequency such that time resolution becomes arbitrarily good at high frequencies while frequency resolution becomes arbitrarily good at low frequencies.

Generally and mathematically, a multi-resolution analysis consists of an increasing sequence of successive approximation spaces V_j, ranging from coarse to fine, which are closed subspaces within and satisfy (following the notation of Daubechies (Ref. 5)

$$\{0\} \rightarrow \cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \cdots \rightarrow L^2(\Re) \quad \Rightarrow NESTED$$

$$\overline{\bigcup_{j \in Z} V_j} = L^2(\Re) \quad \Rightarrow CLOSURE$$

$$\bigcap_{j \in Z} V_j = \{0\} \quad \Rightarrow ORTHOGONALITY$$

$$x(t) \in V_j \iff x(at) \in V_{j+1}, \quad a > 1 \quad \Rightarrow MULTI - RESOLUTION.$$

If P_j is an orthogonal projection operator onto V_j, then the closure property ensures that a suitable approximation to $x(t)$ can be ultimately obtained, i.e.,

$$\lim_{j \to -\infty} P_j x(t) = x(t) \quad \forall x \in V_j.$$

The multi-resolution aspect of these *ladder spaces* is due to the last of the above properties, where all spaces are a *scaled* version of each other such that ultimately they are scaled versions of a central subspace V_0, i.e.,

$$x(t) \in V_j \iff x(a^j t) \in V_0, \quad a > 1.$$

Here the notion of *scale* is that of cartography in that, at a given scale, the signal is replaced by a best approximation that can be *drawn* at that scale. By moving from a coarse scale to a fine scale, one zooms in on the details of the signal, and, hence, a more exact representation of the original

signal. Obviously, the original sampled signal is all the information that exists and this then becomes the finest resolution in information that is possible; i.e., all further processing can only decompose the signal to those bits of information contained in the signal without creating any new information.

Multi-resolution analysis can be thought of as an approximation sequence based upon successive decimation filters where the finest time resolution is the original (typically oversampled) signal. Subsequent coarser time domain approximations can then be obtained by repeated application of decimate-by-n filtering, where every n samples are removed in the time domain. In the frequency domain, this results in peeling away the higher frequencies with each application of the decimation filter, since the Nyquist criterion must be satisfied. Now to increase low-frequency resolution, one simply accumulates more time-domain samples to perform a larger frequency domain transform such that the same N-point FFT results in $N/2$ frequency bins distributed over a much smaller frequency bandwidth due to the filtering and decimation process [Dr. B.W. Bomar, UTSI]. Consequently, time domain resolution is sacrificed at the expense of frequency domain resolution.

Notice that one can also look at multi-resolution analysis as a contraction and expansion (or *dilation*) of the support of the projection/approximation basis functions. This is due to the fact that at high frequency, *best* may be determined to be an N-sampled FFT; therefore, the time analysis window is Nt_0, where t_0 is the sample time interval. This means that the support of the analysis is much narrower for the higher frequencies than for the lower frequencies accumulated over nNt_0, where an n-decimation was used for the same constant $q = \Delta f/f$.

4.2 WAVELET THEORY

Wavelet theory unifies various signal processing techniques developed independently to overcome the problems of Fourier analysis. For example, multi-resolution signal processing, image compression sub-band coding, and wavelet series expansions are all considered a single theory mathematically (Ref. 6). Wavelet theory provides a good technique for signal approximation using scaled basis functions, of which Fourier analysis is just one application (where the support is infinite, non-decaying, and the basis functions are sine functions), and allows one to examine the effectiveness of the approximation on deeper mathematical levels systematically. This paper is primarily concerned with the application of wavelets to non-stationary one-dimensional signals, and therefore will limit its discussion of wavelets to a signal processing point of view. An additional interest is in the use of wavelets to systematically improve the frequency resolution of low-frequency components embedded within broadband signals, whether stationary or non-stationary, without unnecessarily burdening the computation of high-frequency spectral components. This then will allow real-time low-frequency pattern recognition embedded in broadband signals such as required for bearing fault diagnosis based upon engine case accelerometer data (Ref. 7).

The Fourier Transform spanned the space $L_2(0, 2\pi)$ where the elements $x(t)$ satisfy

$$\int_0^{2\pi} |x(t)|^2 dt < \infty .$$

However, for transient signals, there may be no periodicity (at least for the transient component even if embedded in a periodic signal); therefore, we must look for a transform which span $L^2(\Re)$ where the elements $x(t)$ satisfy

$$\int_{-\infty}^{\infty} |x(t)|^2 dt < \infty .$$

As stated by Chui (Ref. 2, p. 3),

> Clearly, the two function spaces [for each of these transforms] $L^2(0, 2\pi)$ and L^2 ($\Re$) are quite different. In particular, since (the local average values of) every function in $L^2(\Re)$ must "decay" to zero at $\pm\infty$, the sinusoidal (wave) functions ω_n do not belong to $L^2(\Re)$. In fact, if we look for "waves" that generate $L^2(\Re)$, these waves should decay to zero at $\pm\infty$; and for all practical purposes, the decay should be very fast. That is, we look for small waves, or "*wavelets*", to generate L^2 ($\Re$).

The rapid decay of the *wavelet* provides the localization necessary to adapt to local transients in a non-stationary waveform. The fact that we look for a wave or oscillating function at all is to somehow preserve the concept of frequency for vibration/harmonic analysis, and we also desire localization in the frequency domain as well.

Now we still want these wavelets to be generated from a single function just as in the Fourier basis functions where

$$w_n(t) = w(nt) = e^{in\omega_o t} \quad \forall\ integers\ n$$

to ease the analysis. In addition, we still want the functions to be complete and orthogonal in $L^2(\Re)$ so we can be assured that the approximation converges to $x(t)$ in the limit and that we can successively find each coefficient of the approximation.

To span L^2 ($\Re$) using a single function which decays rapidly, i.e., is compactly supported on the real line, it is necessary to shift the function along $\Re$. Consider ψ to be a wavelet basis function, real, and of compact support on $\Re$; then for ψ to span all of $\Re$,

$$\psi_b(x) = \psi(x - b) \quad x, b \in \Re.$$

15

This is the translation property of ψ. Now, obviously ψ must still be capable of describing different frequencies, since our intent is to perform a frequency analysis of the mechanical phenomena. To accomplish this easily, without the use of sinusoids, we must consider the fundamental definition of frequency as the inverse of the period between zero crossings or the time extent of the compactly supported analysis functions. With this in mind, we consider frequency as embedded in the dilations of the basis function such that

$$\psi_{a,b}(x) \approx \psi(ax - b) \qquad x, a, b \in \Re, \quad a > 0$$

Observe that our development parallels that for Fourier analysis, and seeks a basis function that is obtained from a single function by dilations (ax) to adjust to varying frequencies and translations (b/a) to span the real line, i.e.,

$$\psi_{a,b}(x) \approx \psi(ax - b) \approx \psi\left[a\left(x - \frac{b}{a}\right)\right] \qquad x, a, b \in \Re, \quad a > 0,$$

and is different from the Fourier analysis only in the basis functions used, which are not periodic and are of compact support.

Now to normalize this basis in $L^2(\Re)$, consider

$$\| \psi(ax - b) \|_2 = \left[\int_{-\infty}^{\infty} |\psi(ax - b)|^2 dx\right]^{\frac{1}{2}} = a^{-\frac{1}{2}} \left[\int_{-\infty}^{\infty} |\psi(x)|^2 dx\right]^{\frac{1}{2}} .$$

Therefore

$$\| \psi(ax - b) \|_2 = a^{-\frac{1}{2}} \| \psi(x) \|_2$$

so that for $\psi_{a,b}(x)$ to be of unit length,

$$\psi_{a,b}(x) = a^{\frac{1}{2}} \psi(ax - b) \qquad x, a, b \in \Re, \quad a > 0 .$$

For economy of the analysis, let's restrict ourselves to integral dilations and translations just as we restrict ourselves to integral frequencies for the FFT. Let Z denote the set of integers; then for $j \in Z$, $a \rightarrow a_j = a_o^j$, $a_o > 1$, where a_o is the fixed dilation parameter. Now since the width (support) of the analyzing wavelet will be proportional to a_o^j, it follows that we will need to translate in steps proportional to the width of the wavelet basis function to cover the whole real line. Consequently, let $b \rightarrow b_k = kb_o a_o^{-j}$, $k \in Z$, where b_o is fixed. Then

$$\psi_{j,k}(x) = a_o^{\frac{j}{2}} \psi(a_o^j(x - kb_o a_o^{-j}))$$

$$= a_o^{\frac{j}{2}} \psi(a_o^j x - kb_o) \qquad x \in \Re, \quad a > 1, \quad b_o > 0, \quad j, k \in Z.$$

16

Later, b_o will become the sampling rate; however, for now let $b_o = 1$, where it is assumed that the shift k is proportional to the width of ψ.

Since our interest is in the analysis of signals evolving in time, first let us look at what the time description of the wavelet is. Now time is continuous just like the real line. However, one still wants to perform integral dilations and translations utilizing the set of integers Z which is a subset of $\Re$. Therefore, the wavelets describing a time evolving signal instead of a spatially evolving one would be [in Daubechies' notation]

$$\psi_{j,k}(t) = a^{-\frac{j}{2}}\psi(a^{-j}t - k\tau) \qquad a > 1, \quad j,k \in Z,$$

and where τ = sample time interval.

In summary, the wavelets are basis functions for the wavelet expansion just as the complex exponentials (sinusoids) are the basis functions for the Fourier expansion. The continuous wavelet expansion of a function (signal) $x(t)$ in L^2 ($\Re$) is the inner product defined as

$$W[x(t)] = (T^{wav}x)(a,\tau) = |a|^{-\frac{1}{2}}\int_{-\infty}^{\infty} x(t)\psi(\frac{t-\tau}{a})dt$$

or discretized[2]

$$W[x(t)] = (T^{wav}_{j,k}x)(a_o,\tau) = |a|^{-\frac{1}{2}}\int_{-\infty}^{\infty} x(t)\psi(a_o^{-j}t - kT)dt \ = \ <x, \psi_{j,k}>$$

where $a \to a_o^j, \tau \to ka_o^jT$, and $a_o > 1$ (generally 2), j, k $\in Z$. The coefficient outside the integral is the normalization factor such that the L^2-norm $\|\psi\|_2 = 1$. The a_o-term is the magnification where j negative and large corresponds to large magnification, and consequently $a_o^jT = T/a_o^{|j|}$ corresponds to small steps, and, therefore, fine details are discernible (i.e., captures high-frequency, short transient characteristics).

The major mathematical difference between wavelets and the Fourier basis functions is the requirement that

$$\int_{-\infty}^{\infty} \psi(t)dt = 0,$$

---

2.   Here *discretized* means that only integral shifts in scale and window position are to be considered. However, this is still the **continuous** wavelet transform (CWT) since the integral and $x(t)$ are continuous in time, and this *discretization* is merely for analysis economy.

which allows the introduction of the dilation (or scale) parameter in order to make the time-frequency window flexible (Ref. 2). However, this also means there is no DC-component in wavelet analyses, but then again, the time-frequency window would have to be infinite to hold the DC-component. What this finally implies is that the final expansion can only reduce the signal to an m-point approximation (for an mth-order wavelet) in time, not a single DC component.

The wavelet is developed not from just a single mother analyzing function as in the Fourier case, but from a single scaling function which is orthogonal to the analyzing wavelet basis function. In contrast to the STFT, which essentially windowed the basis functions (or transforms only) of the expansion, the wavelet truly does window the time domain and the Fourier transform directly. This leads naturally to a filtering/decimation process. Consider $\phi(t)$ to be a smoothing function which removes irregularities by averaging the signal locally over some time (or sample) interval. If one performs the inner product of $x(t)$ and $\psi(t)$, then one obtains an approximation of $x(t)$ within the time resolution of $\psi(t)$, i.e., $< x\psi >$, is a smoothed version of x. The scaling function must also cover the real line and *average* over varying intervals. Consequently, the scaling function has the same form as the wavelet, i.e.,

$$\phi_{j,k}(t) = a^{-\frac{j}{2}}\phi(a^{-j}t - k\tau) \qquad a > 1, \quad j, k \in Z,$$

and where again τ = sample time interval. However, the scaling functions only window; they don't wiggle or oscillate.

The wavelet function extracts the detail from the signal, the difference between the original and smoothed versions, such that the signal approximation is composed of the smoothed version plus the detail removed by the smoothing operation. Therefore, two function spaces are established:

$$\phi \in V \Rightarrow \{0\} \subset \cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \cdots \rightarrow L^2(\Re) \quad \Rightarrow \quad SCALE\ SPACE$$

and

$$\psi \in W \Rightarrow \{0\} \oplus \cdots \oplus W_2 \oplus W_1 \oplus W_0 \oplus W_{-1} \oplus W_{-2} \oplus \cdots \rightarrow L^2(\Re) \quad \Rightarrow \quad DETAIL\ SPACE$$

These spaces are related in that

$$V_{j-1} = V_j \oplus W_j,$$

18

i.e., the next finer level of approximation is generated by the orthogonal sum of the smoothed approximation plus its detail lost due to smoothing at that level. Note that the wavelet spaces W_j inherit the scaling property, i.e., $x(t) \in V_j \Leftrightarrow x(a^j t) \in V_o$, from V_j; therefore, $x(t) \in W_j \Leftrightarrow x(a^j t) \in W_o$. Also note that W_j is the orthogonal complement of V_j in V_{j-1}, i.e., W_j, the detail, does not exist in V_j; it exists in the next finer level of approximation, V_{j-1}. Consequently, for $j < J$,

$$v_j = V_J \oplus \bigoplus_{k=0}^{J-j-1} W_{J-k} , \quad L^2(\Re) = \bigoplus_{k \in Z} W_k ,$$

defines a multi-resolution pyramidal decomposition where all of the spaces are orthogonal, i.e.,

$$V_j \perp V_{j'}, \quad W_j \perp W_{j'}, \quad j \neq j' \quad \text{and} \quad V_j \perp W_j.$$

Note that V_j is decomposed into a coarser approximation plus its lost detail, but **never** do we decompose the detail any further, as shown below. This way the energy of the signal is preserved.
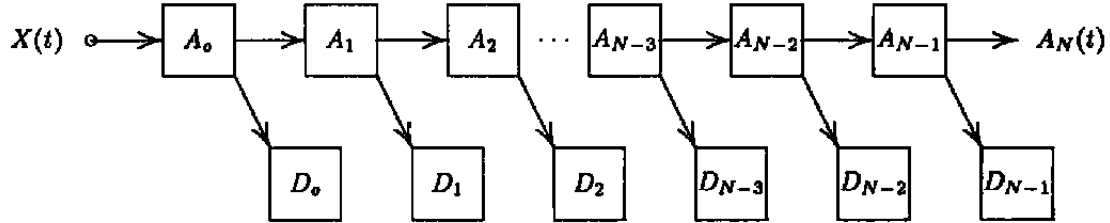
WAVELET PYRAMIDAL DECOMPOSITION

Since the interest in wavelets is for the on-line analysis of a signal which has been digitally sampled, it makes sense to look at wavelets from a filtering (digital signal processing) point of view. Let $a = 2$, then the wavelet will have binary dilations ($2^j n\Delta t$) and dyadic translations ($k\Delta t/2^j$) where

$$\psi_{j,k}(n\Delta t) = 2^{-\frac{j}{2}} \psi(\Delta t(2^{-j} n - k)),$$

or simply,

$$\psi_{j,k}(n) = 2^{-\frac{j}{2}} \psi(2^{-j} n - k).$$

Notice that dilating a signal by a factor of 2 is equivalent to subsampling in the sense that only half the samples are necessary for similar resolution in a given time interval per Nyquist. Also notice that, because of the duality between frequency and time, for a binary dilated signal (1/2 the frequency) twice the time interval is required to sample that signal as compared to the original signal, non-dilated version, for a similar level of resolution. Consequently, one can lowpass filter, determine the detail lost by subtracting the lowpassed version from the original, and then decimate by a factor of 2 and accumulate more samples (look at longer time intervals) to get any resolution in frequency as desired at the expense of time resolution. This is the MRA described above. However, this MRA requires the two basis functions described by the wavelet approach. In this case, ϕ performs the lowpass filtering and decimation to coarsen the approximation (smooth the

signal), and ψ **simultaneously** performs the highpass filtering at the previous level to preserve the detail (fluctuations) which will be lost going to this coarser level of approximation. The wavelet analysis then processes the data according to the grid as shown below.
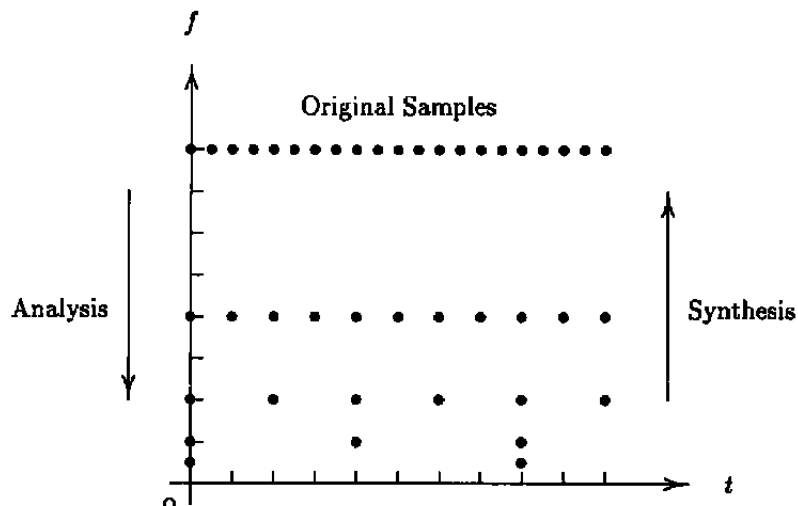
WAVELET CALCULATION GRID

Note again that no DC processing is performed, that the grid is much more sparse than that for the STFT, and finally that not all frequencies are processed; rather, the frequency axis is scaled according to a base 2 logarithm due to the half-band filtering and subsequent decimation by two.

As mentioned above, the wavelet (detail) function spaces and the scale function spaces are related. As filters, this places some unique requirements/restrictions on the coefficients such that *perfect* reconstruction is possible. These restrictions result in a class of filters known as quadrature mirror filters (QMF). Quadrature mirror filters (see schematic) are filters G and H which, for all signals X of finite energy, produce

$$\| Y_L \|^2 + \| Y_H \|^2 = \| X \|^2$$

where Y_L is the output from the lowpass filter G and Y_H is the output from the highpass filter H, both of which are operators which map $l^2(Z) \Rightarrow l^2(2Z)$, where l is the Hilbert space such that $\sum_{n=-\infty}^{\infty} |x(n)|^2 < \infty$ for $n \in Z$. In the case of perfect halfband filters, it is obvious that the coarser approximation Y_L contains only those frequencies below $\pi/2$ of the original signal $x(n)$, while the highpassed signal Y_H contains only those frequencies above $\pi/2$. Consequently, decimation is justified for both the coarser approximation and the detail, holding the number of coefficients ($\{A_1\}$, $\{D_o\}$) constant at the number of original data samples. Note that reconstruction is perfect, and this may be expressed by

20

$$I = G'G + H'H$$

where G' and H' are the adjoints of the filters G and H, respectively, mapping $l^2(2Z) \Rightarrow l^2(Z)$. In addition, note that if $g(n)$ denotes the impulse response of the lowpass filter with frequency response $G(\omega)$, a highpass filter can be obtained by translating $G(\omega)$ by π-radians (i.e., replacing ω by $\omega - \pi$) and correcting for phase. Therefore, the highpass frequency response $H(\omega) = G^*(\omega - \pi)$, or $h(n) = e^{-j\pi n} g(1 - n) = (-1)^n g(1 - n)$, which is true for finite impulse response (FIR) filters. The following block diagram illustrates the quadrature mirror filtering process.
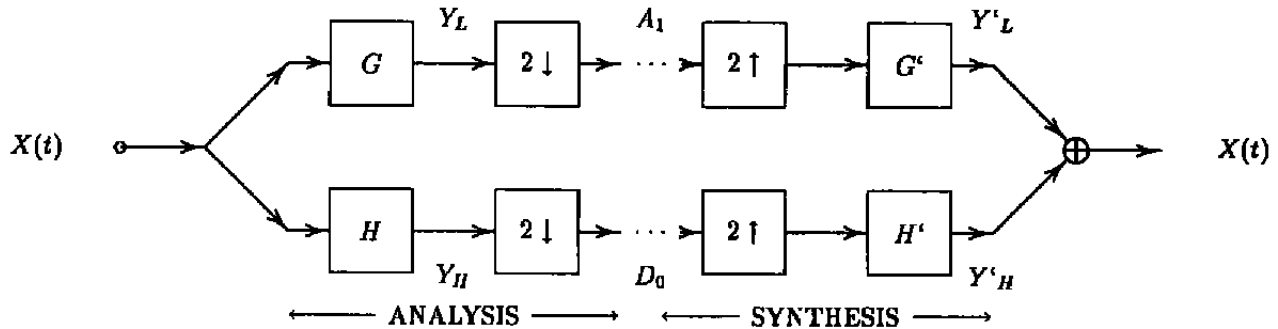
QUADRATURE MIRROR FILTERING

Daubechies (Ref. 5) has shown that such filters can be constructed with all of the desirable characteristics, i.e., compact support in both the frequency and time domains, orthogonality, minimal coefficients (no more than the number of original samples) and perfect reconstruction. In this case the scale functions and wavelet functions, respectively, are

$$\phi_{j,k}(n) = 2^{-j} \sum_k g_k \phi(2^{-j}n - k)$$

and

$$\psi_{j,k}(n) = 2^{-j} \sum_k h_k \psi(2^{-j}n - k),$$

which are known as two scale difference equations or dilation equations, and $h(n) = (-1)^n g(n - 1)$.

The construction of wavelets then begins with the scaling function ϕ. Newland (Ref. 8, pp. 308–321) provides an excellent explanation as to how these filter coefficients are determined based upon:

- **CONSERVATION of AREA** (Unique Scaling Function):

$$\sum_{k=0}^{N-1} g_k = \sqrt{2}$$

- **ACCURACY:**

$$\sum_{k=0}^{N-1} (-1)^k k^m g_k = 0, \qquad m = 0, 1, 2, \ldots, \frac{N}{2} - 1$$

- and **ORTHOGONALITY:**

$$\sum_{k=0}^{N-1} g_k g_{k+2m} = 0, \quad m \neq 0 \qquad and \qquad \sum_{k=0}^{N-1} g_k^2 = 1,$$

where N is the number of coefficients desired. These conditions result in a set of $N + 1$ simultaneous equations in terms of g_k, and finally, $h_k = (-1)^k g_k$. Assuming that such filters are available, the transformation of time domain data to a time-scale representation resulting from convolutions of the signal with these filters follows the pyramidal scheme as coded by Press (Ref. 9) and illustrated below.
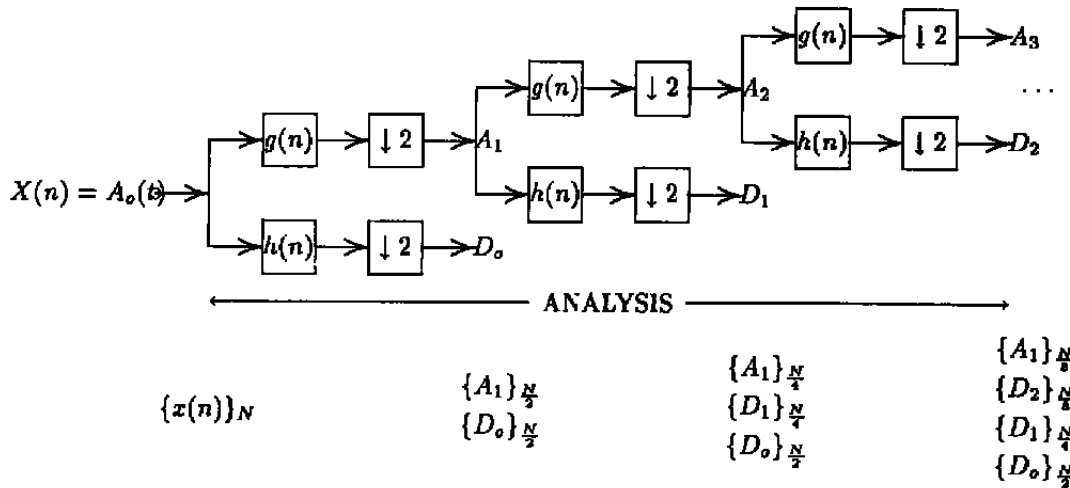
DISCRETE WAVELET TRANSFORM PYRAMIDAL ALGORITHM

The most unique aspect of wavelet analysis is that the time-domain data is transformed not into the familiar 2D amplitude-frequency plane as with Fourier analysis, but rather into a 3D amplitude-scale-time domain. Again, *scale* refers to the broadening of the basis function to fill the time (or space) interval associated with the filter order, i.e., an mth-order wavelet filter requires m data points for the convolution process regardless of the level of the analysis. Consequently, the basis function is more compact for the first level of wavelet analysis and broadens in time (or space)

with each successive level of the analysis due to the decimation process which lengthens the time (or space) interval between successive data points. The inverse of this scale, or *period-of-existence,* of the basis function takes on a similarity to *frequency* in a mechanical sense if the basis function has a single zero-crossing as in the Haar and 4-tap Daubechies wavelets. This is due to the fact that a material fiber would undergo a tensile/compressive cycle at a rate of once per scale of the basis function, similar to that of a sine wave in Fourier analysis.

This can be better shown if one considers a sine wave signal composed of two harmonics plus the fundamental for Fourier analysis sampled at 8 times per cycle for the highest harmonic, and the same except a square wave for Haar wavelet analysis as shown in Fig. 6. The Fourier analysis applied to the sine waves illustrates each of the harmonics directly in the Fourier (amplitude vs. frequency) plane as coefficients of the sine waves making up the signal represented as spikes of the appropriate amplitude at the appropriate frequency. The wavelet analysis applied to the square waves illustrates these harmonics in the 3D time-scale plane as **detail** coefficients, which are constant valued functions of the appropriate amplitude at the appropriate scale of the basis function extending for all time (or space) of the sample block. These functions are constant-valued in time since the square wave signal is periodic, just as the Fourier coefficients are constant in time (or space) since Fourier analysis assumes stationary signals and all temporal (or spatial) dependence is lost following the FFT.

Recall that the wavelet detail coefficients are the amplitudes of the fluctuating *continuous* wavelet basis functions, just as the FFT coefficients are the amplitudes of the *continuous* sine wave basis functions, to be combined to synthesize the respective signals. Similarly, the wavelet scale coefficients are the decimated signal samples, i.e., a smoothed representation of the signal at a coarser scale. For the Haar wavelet applied to a square wave oversampled at 8 samples per cycle, the wavelet amplitudes are all zero except the third level of decomposition (for 512 samples, this is level 6 of the 9 possible, i.e., $2^9 = 512$, level 9 being the original samples), which is constant for all values of time (or space) for the periodic square wave. Consequently, the Haar wavelet does for square waves what Fourier analysis does for sine waves, i.e., determines the exact scale or frequency match, respectively, for a signal composed of the respective basis functions. The Haar wavelets, of course, can only provide such perfect scale matches if the data is properly aligned to harmonics and sampled appropriately, but this, too, is no different than the FFT, where pure spikes will occur only if the data has integral frequencies and is sampled with zero phase. Otherwise, smearing occurs in the FFT spectrum, and the wavelet analog is to have numerous non-zero, and typically oscillatory, details at scales **below** the highest harmonic in the signal. It should be noted that at the third level of decomposition, the wavelet analysis has decimated the original signal by a factor of 8, the exact factor of the samples per cycle, leaving only **one sample per cycle.** This holds for other sampling factors as well. In other words, the log_2 of the number of samples per cycle is the

decomposition level which will be non-zero if the signal is a series[3] of the basis function of the wavelet. At this level, the reciprocal of the interval between the remaining samples is the *frequency* of the square wave.

It should also be pointed out that the choice of the wavelet basis function is equivalent to making an assumption about the composition of the signal. The use of a Haar wavelet for a square wave is making the obvious assumption that a square wave is made up of the Haar basis, just as in the case of using the FFT makes the assumption that the signal is composed of sine functions. Each basis applied to the wrong signal will generate broad spectra which tend to yield little or no information about the signal composition. The potential of wavelets is that we now have a broad selection of basis functions to apply to a problem to detect characteristics which are smeared utilizing only the FFT, or to perform pattern recognition analysis where we generate a new basis dependent upon the pattern of interest.

4.3 APPLICATION TO MODEL SIGNAL

Consider the same signal used for the STFT analysis discussion of Section 3.0 (see Fig. 1). Application of the Daubechies four-term wavelet (Daub4) to this signal results in Fig. 7. Again, the original signal (wavelet level 10 since the number of samples $= 1024 = 2^{10}$) is shown at scale index 2 and the first wavelet (level 9) is shown at scale index 3. Therefore, subtracting 2 from the scale index in these 3D wavelet plots will yield the wavelet iteration number, which, when subtracted from 10 (the highest wavelet level for these data) yields the wavelet scale. Now the signal has three distinct frequencies, $f_0 = 10$ Hz, $f_1 = 80$ Hz, and $f_2 = 128$ Hz. The sampling rate was $2.56 * F_{max}$ or 327.68 Hz. The first wavelet filters the signal and subsamples by a factor of 2, which means at level 9 the sample rate of the wavelet scale is effectively 163.84 Hz. This is essentially twice f_1, and therefore the 80-Hz signal is shown as a sinusoid at scale index 4, which is level 8 (this was verified using a single 80-Hz component sampled at 128/80 ∙ 2.56 or 4.096 samples per cycle). The trace at scale index 3 (level 9) should be the f_0-modulated 128-Hz cosine wave. Recall, the wavelet shows the detail lost by the filtering at the previous level. The interpretation of this trace is not fully understood by the authors and it is suspected that this trace is noisy due to the wavelet basis/signal mismatch. The Gaussian impulses are observed as spikes at sample time indices 320 and 576 which broaden with successive wavelet iterations. Note that the temporal characteristics were preserved with the wavelet transform, unlike the FFT, and the spikes occur at the peaks of the impulse even though the impulses were rather broad (64 samples each). This is due to the slope change from one side of the peak to the other, and therefore an edge is detected by the wavelet at this point. The trace at scale index 6 (level 6) is an artifact of sampling the 10-Hz signal at a factor of 32.768 (2.56)

---

3. Here series means a series in time (or space) of wavelet basis functions laid end to end.

single 10-Hz component sampled at 32.768 samples per cycle). This artifact appears to be due to a beating between the wavelet zeros and the signal zeros at this level, but this could not be verified by the author. This artifact persists regardless of the order of the wavelet, from Haar to Daub20. The only significant changes going to other wavelet orders were the coarsening of the traces with the Haar wavelet and the smoothing of the traces with the Daub20 wavelet, and the more pronounced spikes at the Gaussian impulse positions due to less noise with the Daub20. The Daub20 wavelet results are shown in Fig. 8. Note that neither Fig. 7 nor Fig. 8 contains any information about the amplitude modulation aspect of the signal. This is a deficiency of the wavelet analysis, at least as it is presently understood and interpreted by the authors.

The STFT provided results which were considered more familiar to that of the wavelet analysis in interpretation aspects, but the wavelet analysis is much faster, has less impact on the hardware requirements for sampling, and provides suitable information once the interpretation is worked out. In addition, the wavelet could be used in multi-resolution schemes to properly filter the signal for FFT analysis or to complement the FFT by searching for impulses or other discrete non-stationary events. Further research into the interpretation of wavelet data must be provided, and perhaps the new harmonic wavelet of Newland (Ref. 8) will shed some light in this area.

4.4 APPLICATION TO PULSE-ECHO ANALYSIS

Signal discontinuities are prevalent in real signals, and one such signal which relies on discontinuities is that of Pulse-Echo, which is a technique for measuring distances in materials nonintrusively through the use of sound waves. In brief, a sound impulse is propagated into the material through an appropriate acoustic coupling and a sensor detects the echo *or reflected wave* caused by the abrupt change in the index of refraction at the opposite edge of the material (this can be the interface between two different materials or an air/material interface). The thickness d of the material is then found as

$$d = c\frac{t}{2}$$

where c is the characteristic speed of sound for the material, and t is the round trip time from initiation of the pulse through the material to the discontinuity in the index of refraction and back through the material to the sensor. The sensors are typically piezoelectric crystals of appropriate natural frequency which act as the transmitter when electrically pulsed and act as the receiver when stimulated by the returning sound wave echo (hence, the name Pulse-Echo). The signals resemble an impulse with ringing, (similar to a damaged bearing signal (Ref. 4) and were simulated by single sided sinc functions with a reduced amplitude for the echo to simulate material transmittance properties. The functions used were of the form

$$x[t] = u_{-1}(t - \tau_1)A_1\frac{sin(t - \tau_1)}{(t - \tau_1)} + u_{-1}(t - \tau_2)A_2\frac{sin(t - \tau_2)}{(t - \tau_2)} \ .$$

where τ_1 was near the beginning of the ensemble, τ_2 was at the midpoint of the ensemble, and A_2 was $0.5A_1$. The first sinc function was spread over the whole interval once started, while the second sinc function started such as to peak at the exact midpoint and build upon the first impulse, hence the unit step functions $u_{-1}(t - \tau)$. The signal was sampled at 16 samples per cycle for the 1-Hz sine wave component of the sinc functions, and 1024 point ensembles were used. Random noise was added to the signal, and various noise floor amplitudes were evaluated. The closeness in time between the simulated pulses could be controlled to simulate varying material thickness. The test parameters are summarized in Table 1, and the wavelet 3D detail plots are shown in Fig. 9.

In pulse-echo measurements, assuming the material speed of sound characteristics are well known, the most important parameter is an accurate measurement between the main pulse and it's echo. Any point on the pulses can be used as a reference point, since the same sensor is used for transmitting and receiving it will have the same response function, but it is necessary to be consistent. Since real signals generally have a significant noise floor, it is usually most accurate to use the peaks in the pulses as the reference points. However, most pulse-echo instruments use a threshold test and capitalize on the rapid decay of the response because detecting the true peak location is difficult electronically. The use of wavelets may change this situation due to their ability to detect edges and rapid slope changes in a signal.

Table 1. Pulse-Echo Test Cases

A_1	A_{NOISE}	N_{START}	Figure
1.0	0.0	128.0	9a
1.0	$0.5A_1$	128.0	9b
1.0	$0.9A_1$	128.0	9c
1.0	0.0	169.0	12a
1.0	0.0	385.0	12a

In pulse-echo measurements, assuming the material speed of sound characteristics are well known, the most important parameter is an accurate measurement of the time interval between the main pulse and its echo. Any point on the pulses can be used as a reference point, since the same sensor is used for transmitting and receiving it will have the same response function, but it is necessary to be consistent. Since real signals generally have a significant noise floor, it is usually most accurate to use the peaks in the pulses as the reference points. However, most pulse-echo instruments use a threshold test and capitalize on the rapid decay of the response because detecting the true peak location is difficult electronically. The use of wavelets may change this situation due to their ability to detect edges and rapid slope changes in a signal.

Various Daubechies wavelets were applied to the signal with the intent of locating the peaks as consistent reference points for determining the time interval. The Haar (Daub2) wavelet provided the best results, and was used for this analysis. The FFT applied to the signal of Fig. 9a merely shows a pulse near DC and offers no timing information (see Fig. 10). It does illustrate that the signal is appropriately sampled, however, since the spectrum is essentially zero well before the Nyquist foldover frequency. Figure 11 illustrates the ability of the wavelet to detect the location of the peaks at wavelet levels 4 and 3. Level 4 is significant since at this level the signal is sampled at essentially the Nyquist criteria. Even in the presence of significant noise (Fig. 9c), the Haar wavelet was able to detect the peaks in either of these two levels. While the fourth level has a second peak which becomes more pronounced with increasing noise, the third level clearly distinguishes which peak in level 4 is the appropriate peak of the pulse for timing purposes. However, the ability of wavelets to determine the peak locations is tied to whether or not the peak was sampled and if the sample index n is a factor of 2^j. This is illustrated in Fig. 12a, where the starting point was at $n = 169$, and a maximum error of 13 percent (for level 3; level 4 is 7 percent) is incurred due to the subsampling resulting in the peak being missed slightly. This is equivalent to the scalloping effect of the FFT when the signal frequency does not correspond to the FFT bin center frequency but rather is offset from the center of the bin. Figure 12b takes this situation to the extreme by narrowing the distances between the pulses and having n a non-integral power of two. This is probably the minimum spacing that would be considered a typical pulse-echo signal (for thinner materials, higher frequencies (tighter pulses) are typically used to space out the pulses for better time resolution of the peaks). The peaks are still well defined in level 4, while level 3 has rounded and smeared the peaks.

In summary, the wavelet analysis provided a consistent set of timing points for improving the resolution of the pulse-echo measurement technique, especially when the signal is buried in significant noise. The limitation of the approach appears to be the pulse peak location relative to an integral power of two sample points within the ensemble. This could be resolved by performing a sliding wavelet analysis such that the peak would become more pronounced as it fell on an integral power of two sample, and then timing between two different wavelets at the same level. Work is being performed to make the sliding wavelet computationally efficient, and to calculate a particular level without having to calculate all previous levels serially.

5.0 MALVAR WAVELET TRANSFORM

Traditional signal analysis using the Fourier transform relies heavily on the fact that the data under study is stationary in some sense. The usual display of the spectra uses a single window and provides no way to show variations in the spectra within the window; usually only one window size is used throughout the analysis, tacitly assuming stationarity at least on the order of several window sizes. Window effects such as leakage are controlled by reducing window influence at the ends of the window interval, thus also discarding much of the influence of the signal data at these transition points and making reconstruction problematic. Noise reduction techniques are based on averaging over several windows which are considered to be stationary in a statistical sense. Signals representing non-stationary phenomena must be treated so as not to discard any parts since it cannot

be assumed that "one part is much like another." This will require basis functions which vary in time locality as well as frequency range. These basis functions both within and between windows must allow complete and stable reconstruction from the analysis coefficients.

Malvar wavelets are a natural development from windowed Fourier transforms. The signal $s(t)$ is first sliced up in time using window functions $w_j(t)$ to obtain a series of windowed signals $w_j(t)s(t)$. Each of the windowed functions is then subjected to Fourier analysis. This closely resembles the Gabor (Ref. 10) approach, since the Fourier transform leads to calculating $fs < s$, $w_{k,j}$ >with the wavelet $w_{k,j}(t) = e_{i\omega kt}w_j(t)$. The windows $w_j(t)$ are chosen to give several advantages.

Gabor used a single Gaussian with offsets for $w_j(t)$ which has a number of disadvantages. The Gaussian does not have compact support and the algorithms must be designed with error analysis in mind. But more serious is the problem of accounting for the overlap of the windows; this becomes clear if we try to reconstruct $s(t)$ from the spectra. The phenomena known as "leakage" in Fourier analysis are caused by windowing which does not completely include a feature (such as a complete cycle) in a single window, but instead divides the influence across two adjacent windows. What is needed for reconstruction (and, as we will see, for control of some other influences) is that we have orthogonality between the windows as well as within them.

Malvar studied lapped Fourier transforms (Ref. 11), which can be regarded as overlapping rectangular windows, with the intent of bringing about orthogonality between windows. Actually, rectangular non-overlapping windows are already orthogonal. Newland exploited this fact to use the Fourier transform of a rectangular window to construct a family of wavelets based on the sync function (Ref. 8). But the edge effects of the rectangular window are severe. Figure 13 shows a reconstruction of an acoustic signal from spectra given in several adjacent windows after discarding several of the least significant spectral coefficients; note the cusp at several interval boundaries caused by the slow roll-off of the Fourier spectrum of the rectangular window. Malvar found orthogonality conditions on the lapped transforms which can be described similarly to a windowed Fourier transform. But his window extends outside the "primary" interval and overlaps the adjacent window. Figure 14 shows how the windows of two adjacent intervals are related while Fig. 15 shows what a typical Fourier basis function looks like after being windowed. The smooth edges on the windows are reminiscent of window "carpentry" which has been used to reduce leakage and other window effects (Ref. 12); but the older techniques did not overlap the windows nor account for inter-window signal effects.

Coifman and Meyer (Ref. 13) modified Malvar's approach to obtain variable width windows while continuing to maintain orthogonality within and between the windows. Their approach still retains the faster roll-off in window effects obtained by using smooth window edges.

Only the window shapes of part of a possible basis are shown in Fig. 16. This basis would be suitable to represent a signal which was almost stationary in region 1 and 4 while significant transient behavior is captured in regions 2 and 3. We must emphasize again that the transform

transient behavior is captured in regions 2 and 3. We must emphasize again that the transform coefficients on each of the intervals do not interact; the orthogonality ensures this. Furthermore, the effect of a feature extending between two adjacent windows is captured such that it can be accurately reconstructed in spite of round-off or truncation errors.

We illustrate the construction of the Meyer-Coifman-Malvar discrete transformation using a cosine type IV basis on each basic window interval. Other bases, including the Fourier exponential basis, can be used, but the symmetries needed to ensure the behavior referred to in the previous paragraph are only simple in the cosine type IV case.

Figure 17 shows the nomenclature for both the continuous and the discrete versions of the Malvar wavelets. The c_j represent the "classic" interval end-points for non-overlapped windows. The jth interval extends from c_j to c_{j+1} and has length l_j.

The window $w^j(t)$ corresponding to the jth interval is constructed by modifying the rectangular window using the bell profile function $\beta^j(t)$ shown in Fig. 18. The result is that the window $w^j(t)$ extends α_j to the left of c_j and α_{j+1} to the right of c_{j+1}. The window is also modified within the original basic interval as shown in the Fig. 18. The "mirror image" of $\beta^j(t)$ is used to modify the right end of the $(j-1)$th interval. The particular $\beta^j(t)$ of Fig. 18 ensures $w^j(t)$ has continuous derivatives even at $c_j \pm \alpha_j$. $\beta^j(t)$ can be chosen to make $w^j(t)$ as smooth as desired.

Discretization is simplified by rescaling the entire problem, if necessary, so that the sample points on the jth interval occur at integers $a_j + n$, $0 \le n < l_j$ and the c_j occur exactly halfway between a_{j-1} and a_j. R_j is the largest integer less than $a_j + 1/2$.

$\beta^j(t)$ is discretized so that $\beta_n^j = \beta^j(a_j + n)$ and other functions of t are similarly characterized.

The discrete window function w_n^j for the jth interval which results is given by

$$w_n^j = \begin{cases} \beta_n^j & \text{if } n \in [-R_j, R_j - 1], \\ 1 & \text{if } n \in [R_j, l_j - 1 - R_{j+1}], \\ \beta_{l_j-n-1}^{j+1} & \text{if } n \in [l_j - R_{j+1}, l_j + R_{j+1} - 1], \\ 0 & \text{otherwise.} \end{cases}$$

The basis functions for conventional discrete cosine type IV transforms are given in continuous and discrete form by

$$\varphi_k^j(t) = \sqrt{\frac{2}{l_j}} \cos[(k + \frac{1}{2})(t - c_j)\frac{\pi}{l_j}], \qquad j \in Z, k \in N, 0 \le t \le l_j$$

$$\varphi_{k,n}^j = \sqrt{\frac{2}{l_j}} \cos[(k + \frac{1}{2})(n + \frac{1}{2})\frac{\pi}{l_j}], \qquad j \in Z, \ 0 \le k, n < l_j$$

Thus, the continuous and discrete basis function are

$$\psi_k^j(t) = w^j(t)\varphi_k^j(t)$$

$$\psi_{k,n}^j = w_n^j \varphi_{k,n}^j$$

This definition does not yet define $\psi_{k,n}^j$ off the interval from c_j to c_{j+1}, but the symmetry of the basis function $\varphi j_k^j(t)$ at the end points and a symmetry requirement on the $\beta^j(t)$ completes the definition.

$$\varphi_{k,-n-1}^j = \varphi_{k,n}^j; \qquad \varphi_{k,2l_j+n}^j = -\varphi_{k,n}^j$$

$$(\beta_n^j)^2 + (\beta_{-n-1}^j)^2 = 1 \qquad j \in Z, \; n \in [-R_j, R_j - 1]$$

The symmetry in β_n^j above is exactly equivalent to the conditions obtained by Malvar which ensure orthogonality between basis functions in two adjacent windows. The orthogonality of the discrete basis functions within and between windows can be proved exactly.

The discrete basis function is given directly in terms of β_n^j by

$$\psi_{k,n}^j = \begin{cases} \beta_n^j \varphi_{k,n}^j & \text{if } n \in [-R_j, R_j - 1], \\ \varphi_{k,n}^j & \text{if } n \in [R_j, l_j - R_j - 1], \\ \beta_{l_j-n-1}^j \varphi_{k,n}^j, & \text{if } n \in [l_j - R_j, l_j + R_j - 1]. \end{cases}$$

Now in order to perform the transform of a signal s_n, we need the transform coefficients $S_k^j = \left(S_{a_{j+n}}, \psi_{k,n}^j \right)$. We will express the scalar product in terms of a new signal $\hat{s}_n$ as follows:

$$S_k^j = \langle s_{a_j+n}, \psi_{k,n}^j \rangle$$

$$= \sum_{n=-R_j}^{l_j+R_j-1} s_{a_j+n} \psi_{k,n}^j$$

$$= \sum_{n=0}^{R_j-1} (s_{a_j+n}\beta_n^j + s_{a_j-n-1}\beta_{-n-1}^j)\varphi_{k,n}^j$$

$$+ \sum_{n=R_j}^{l_j-R_{j+1}-1} s_{a_j+n}\varphi_{k,n}^j$$

$$+ \sum_{n=l_j-R_{j+1}}^{l_j-1} (s_{a_j+n}\beta_{l_j-n-1}^{j+1} - s_{a_j+2l_j-n-1}\beta_{n-l_j}^{j+1})\varphi_{k,n}^j$$

$$= \sum_{n=0}^{l_j-1} \hat{s}_{a_j+n}\varphi_{k,n}^j$$

30

Thus, using the new signal $\hat{s}_n$ the transform becomes

$$S_k^j = \langle \hat{s}_{a_j+n}, \varphi_{k,n}^j \rangle$$

which is an ordinary discrete cosine transform on basic interval j. The special signal $\hat{s}_n$ is said to be obtained by "folding" the signal s_n at c_j and c_{j+1}. As expected, the transform for one window includes the influence of part of the signal in the adjacent windows; it obtains a weighted average behavior at the end-points which smooths the transition across the window edges. A similar "unfolding" rule is obtained easily and blends together the transforms on two adjacent windows to reconstruct the signal without artifacts.

In summary: To perform the Malvar transform, fold at each a_j:

$$\hat{s}_{a_j+n} = \begin{cases} s_{a_j+n}\beta_n^j + s_{a_j-n-1}\beta_{-n-1}^j & \text{if } n \in [0, R_j - 1], \\ s_{a_j+n}\beta_{-n-1}^j - s_{a_j-n-1}\beta_n^j & \text{if } n \in [-R_j, -1]. \\ s_{a_j+n} & \text{otherwise.} \end{cases}$$

then Discrete Cosine Transform $\hat{s}_n^j$ on each interval $[a_j, a_j + l_j - 1]$.

To invert the Malvar transform, first DCT as before (the DCT-IV being used here is its own inverse), then unfold at each a_j:

$$s_{a_j+n} = \begin{cases} \hat{s}_{a_j+n}\beta_n^j - \hat{s}_{a_j-n-1}\beta_{-n-1}^j & \text{if } n \in [0, R_j - 1], \\ \hat{s}_{a_j-n-1}\beta_n^j + \hat{s}_{a_j+n}\beta_{-n-1}^j & \text{if } n \in [-R_j, -1]. \\ \hat{s}_{a_j+n} & \text{otherwise.} \end{cases}$$

5.1 THE WAVELET LIBRARY

Since we may use windows of rather arbitrary size in the signal, it is clear that several choices of the basis intervals and, consequently, of the basis itself are possible. The collection of all the Malvar wavelets on all the intervals in one arrangement of basis intervals is a basis for the L^2 functions on the real line. But there are many choices for the basis intervals, so there are many basis sets for L^2. The collection of all these "bases" is called a wavelet library. The question is: Which basis from the library is the best one?

To see how to compare two bases from the library, we consider the simple case represented in Fig. 19, where we have drawn only the windows for one basis rising above the axis and a window for the second basis but reflected below the axis. If two bases differ only in having the two upper

31

windows in one replaced by the lower window in the other, then we need only decide whether it is better to use the single window (thereby assuring the signal is nearly stationary) or to use the two smaller windows (thereby more accurately capturing any transient behavior).

Wesfried and Wickerhauser (Ref. 14) have applied this idea in acoustic signal processing to recognize speech signals. They measured information in transform for the two intervals together against that in the single interval transform with the formula for "spectral entropy" given by

$$H(S_k^j) = -\sum p_k log p_k \qquad \text{where } p_k = (S_k^j)^2/\|S_k^j\|^2.$$

This is one of several information measures which can be used to "split or merge" the windows to obtain a minimal entropy (maximal information) basis for the signal.

An effective and fast implementation of the idea is to first obtain Malvar transforms on rather small windows, say $l_j = 32$. Then measure the entropy on each pair of intervals separately and grouped as one with $l_j = 64$. If the entropy decreases, keep the larger and discard the smaller. Continue grouping the larger intervals in pairs and replacing if the entropy decreases. The entropy measure and others which are suitable will cause this process to obtain the optimal basis in time comparable to that required by the discrete cosine transform. Also the transform coefficients for the two smaller windows are used recursively to efficiently compute the transform in the larger window.

Figure 20 shows a speech signal reconstructed from an optimized basis. The vertical lines represent the optimal windows for minimum entropy. Note that the more nearly periodic the signal is, the longer the window. This is exactly the behavior we hope to establish for suitable optimization of engine bearing failure data.

5.2 APPLICATION TO BEARING FAULT DATA

Techniques to detect and analyze the approach of bearing failure in rotating machinery have obvious economic benefits. In Fig. 21 are shown the typical Fourier spectra of four types of bearing failure. The spectra are, of course, obtained after failure has led to the reestablishment of stationarity.

The signal for a typical impending failure is shown in Fig. 22. Note the long interval of stationarity before the failure, the detailed non-stationary interval at failure onset, and the following stationary interval which we call "stationary failure."

Fourier analysis of the data using 1024 point transforms is shown in Fig. 23. The spectra give accurate information for the two stationary regimes, but are not useful in the transient regimes.

32

The Malvar transform of the signal is shown in Fig. 24 using 1024 point transforms. Even though this transform is cosine based, we see the qualitative similarity for the three regimes. Indeed, spectra for the stationary regimes are readily recognized.

Clearly, a long Malvar window can replace the shorter windows in the stationary regimes. In the transient regime, on the other hand, much narrower windows are needed to display significant details.

Future work on this signal will include the implementation of the basis optimization algorithm described in the previous section so that the splitting and merging of the windows can be done automatically and, at some point, in real time.

6.0 CONCLUSIONS

Wavelet theory has been presented and contrasted against Fourier analysis from a basis approach for Malvar and Daubechies wavelets. The wavelet analysis was shown to perform similarly to that of Fourier analysis while maintaining the element of time in the transformation. The theory is still new and requires additional work to resolve interpretation issues, but it has been shown to be a valuable analysis tool for transient data analysis, especially if combined with traditional Fourier analysis where appropriate. One area where wavelets offer significant analysis improvement is in the detection of transient edges, and further work continues in this area for image processing and other applications. New wavelets such as the Malvar and Harmonic wavelets are becoming available which have promise for improving the application and interpretation of wavelet analysis.

The Malvar wavelets were chosen for engine bearing failure analysis because their structure allows representation of stationary signals typical of rotating machinery as well as representation of the transient effects that mark the onset of bearing failure. Their similarity to the Fourier transform makes them familiar to engineers already accustomed to using Fourier spectra to understand engine behavior. The use of on-line window length optimization will focus attention on the unusual behavior of an engine, while more normal stationary behavior can be quickly grasped and assimilated. Very efficient lossless data compression is an immediate benefit of the optimization; it also generates effective pattern recognition parameters for use by higher level classifiers such as neural networks.

The analysis of actual engine data indicates more than fair likelihood that the Malvar approach will, in fact, deliver the benefits to analysis for which they were chosen. All but the optimization has been completed with algorithms which are highly efficient and readily ported to digital signal processing computers; these algorithms are also parallelizable and vectorizable. Optimization can

be done in, at worst $O(n \log(n))$ time, but final determination has not yet been made of the actual speed, or of its suitability for parallelization.

Our studies of both pulse-echo data and bearing failure detection point up some important facts concerning wavelet analysis. First, wavelet analysis includes a wide range of alternatives and supplements to Fourier analysis, some already well known and some new. As such, wavelets do not replace Fourier analysis; they broaden it. Second, the wavelet technique used for a problem must be chosen appropriate to the problem if improvements to Fourier analysis are to be realized. Furthermore, where Fourier analysis already provides well-understood insight, wavelet techniques should complement the Fourier technique. Third, the evaluation and use of wavelet techniques requires a thorough foundation in Fourier analysis as well as additional training in a variety of wavelet analysis techniques; while the results of well-designed wavelet analysis are easily comprehended, casual acquaintance with the techniques is simply inadequate for the judgment required to produce a successful analysis design.

REFERENCES

1. Proakis, J. G. and Manolakis, D. G. *Introduction to Digital Signal Processing*. Macmillan Publishing Company, New York, 1988.

2. Chui, C. K. *An Introduction to Wavelets*. Academic Press, Inc., San Diego, CA, 1992.

3. Strang, G. *Introduction to Applied Mathematics*. Wellesley-Cambridge Press, Wellesley, MA, 1986.

4. Li, C. J. and Ma, J. "Bearing Localized Defect Detection Through Wavelet Decomposition of Vibrations," *Economic Implications of Mechanical Failure Prevention*, H. C. Pusey and S. C. Pusey, eds., Vibration Institute, Willowbrook, IL, 1992, pp. 53-62.

5. Daubechies, I. "Ten Lectures on Wavelets," Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

6. Rioul, O. and Vetterli, M. "Wavelets and Signal Processing." *IEEE Signal Processing Magazine*, October 1991, pp. 14–38.

7. Hite, S. W., III. "An Algorithm for Determination of Bearing Health through Automated Vibration Monitoring." AEDC-TR-93-19 (AD-A274 591), December 1993.

8.   Newland, D. E. *An Introduction to Random Vibrations, Spectral, & Wavelet Analysis.* Longman Scientific & Technical, John Wiley & Sons, New York, 1993.

9.   Press, W. H. *Wavelet Transforms.* Harvard-Smithsonian Center for Astrophysics, Preprint, Cambridge, MA, 1991.

10.   Gabor, D. "Theory of Communication," *Journal of the Institute of Electrical Engineering,* Vol. 93, No. III, London, 1946, pp. 429-457.

11.   Malvar, H. S. *Signal Processing with Lapped Transforms.* Artech House, Inc., Norwood, MA, 1992.

12.   Ramirez, R. W. *The FFT, Fundamentals and Concepts.* Prentice-Hall, Inc., Englewood Cliffs, NJ, 1985.

13.   Coifman, R. R., Meyer, Y., and Wickerhauser, V. "Wavelet Analysis and Signal Processing," *Wavelets and Their Applications,* Mary Beth Ruskai, et al., eds., Jones and Bartlett Publishers, Boston, MA, 1992.

14.   Wesfreid, E. and Wickerhauser, M.V. "Adapted Local Trigonometric Transforms and Speech Processing," *Transactions on Signal Processing,* Vol. 41, No. 12, December 1993, pp. 3596-3600.

BIBLIOGRAPHY

Bendat, J. S. and Piersol, A. G. *Engineering Applications of Correlation and Spectral Analysis.* John Wiley & Sons, New York, 1980.

"The Cepstrum: A Guide to Processing," *Proceedings of the IEEE,* Vol. 65, No. 10, October 1977, pp. 1428–1443.

Chui, C. K. *An Introduction to Wavelets.* Academic Press, Inc., San Diego, CA, 1992.

Cody, M. A. "The Fast Wavelet Transform," *Dr. Dobb's Journal,* April 1992, pp. 16–28.

Coifman, R. R., Meyer, Y., and Wickerhauser, V. "Wavelet Analysis and Signal Processing," *Wavelets and Their Applications,* Mary Beth Ruskai, et al., eds., Jones and Bartlett Publishers, Boston, MA, 1992.

Courrech, J. "Condition Monitoring of Machinery," *Shock and Vibration Handbook*, edited by C. M. Harris, McGraw-Hill Book Company, New York, 1988, pp. 16–1– 16–18.

Curtis, A. J. "Concepts in Vibration Data Analysis," *Shock and Vibration Handbook*, edited by C. M. Harris, McGraw-Hill Book Company, New York, 1988, pp. 22–1 — 22–28.

Daubechies, I. "Ten Lectures on Wavelets," Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

Daubechies, I. "The Wavelet Transform, Time-Frequency Localization and Signal Analysis," *IEEE Transactions on Information Theory*, Vol. 36, No. 5, September 1990, pp. 961–1005.

Janse, C. P. and Kaizer, J. M. ""Time-Frequency Distributions of Loudspeakers: The Application of the Wigner Distribution," *Journal of Audio Engineering Society*, Vol. 31, No. 4, April 1983, pp. 198–223.

Joen, J. J. and Shin, Y. S. "Pseudo Wigner-Ville Distribution, Computer Program and its Applications to Time-Frequency Domain Problems," Naval Postgraduate School Report NPS-ME-93-002, March 1, 1993.

Li, C. J. and Ma, J. "Bearing Localized Defect Detection Through Wavelet Decomposition of Vibrations," *Economic Implications of Mechanical Failure Prevention*, edited by H. C. Pusey and S. C. Pusey, Vibration Institute, Willowbrook, IL, 1992, pp. 53–62. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, July 1989, pp. 674–693.

Meyer, Y. (Translated by Ryan, R.D.) *Wavelets: Algorithms and Applications*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1993.

Najafi, B. and Hakim, H. "A Comparative Study of Non-Parametric Spectral Estimators for Application in Machine Vibration Analysis," *Machine Systems and Signal Processing*, Vol. 6, No. 6, Academic Press, London, 1992, pp. 551–574.

Permann, D. and Hamilton, I. "Wavelet Analysis of Time Series for the Duffing Oscillator: The Detection of Order Within Chaos," *Physical Review Letters*, Vol. 69, No. 18, November 1992, pp. 2607–2610.

Proakis, J. G. and Manolakis, D. G. *Introduction to Digital Signal Processing*. Macmillan Publishing Company, New York, 1988.

Ramirez, R.W. *The FFT, Fundamentals and Concepts.* Prentice-Hall, Inc., Englewood Cliffs, NJ, 1985.

Resnikoff, H. L. "Wavelets and Adaptive Signal Processing," *Optical Engineering,* Vol. 31, No. 6, June 1992, pp. 1229–1234.

Resnikoff, H. L. and Burrus, C. S. "Relationships Between the Fourier Transform and the Wavelet Transform," SPIE, Vol. 1348, *Advanced Signal-Processing Algorithms, Architectures, and Implementations,* July 1990, pp. 291–300.

Rioul, O. and Vetterli, M. "Wavelets and Signal Processing," *IEEE Signal Processing Magazine,* October 1991, pp. 14–38.

Strang, G. "Wavelet Transforms versus Fourier Transforms," *Bulletin of the American Mathematical Society,* Vol. 28, No. 2, April 1993, pp. 228–305.

Strang, G. "Wavelets and Dilation Equations: A Brief Introduction," *SIAM Review,* Vol. 31, December 1989, pp. 614–627.

Tucker, Michael. "Adaptive Wavelet Signal Processing," (AD-B166 456), U.S. Army Strategic Defense Command, Huntsville, AL, July 1992.

Wesfreid, Eva and Wickerhauser, M. V. "Adapted Local Trigonometric Transforms and Speech Processing," *Transactions on Signal Processing,* Vol. 41, No. 12, December 1993, pp. 3596–3600.
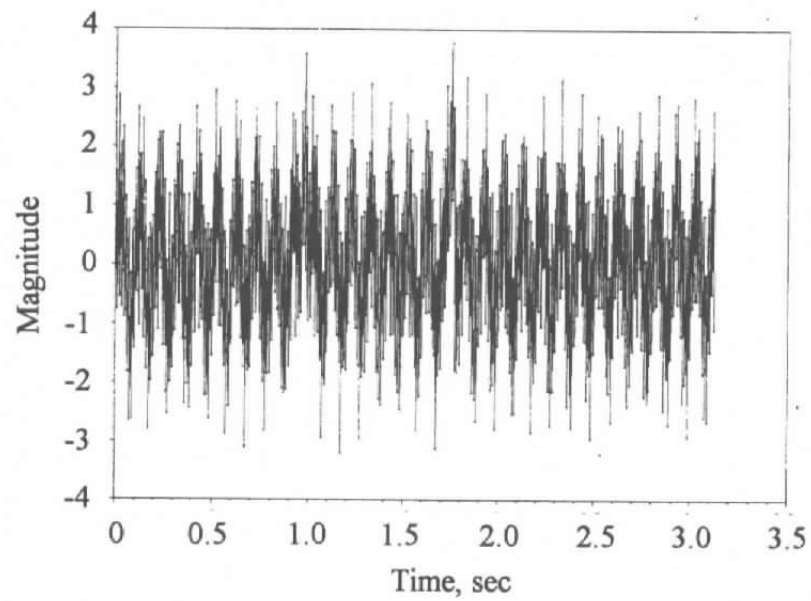
a. Example signal
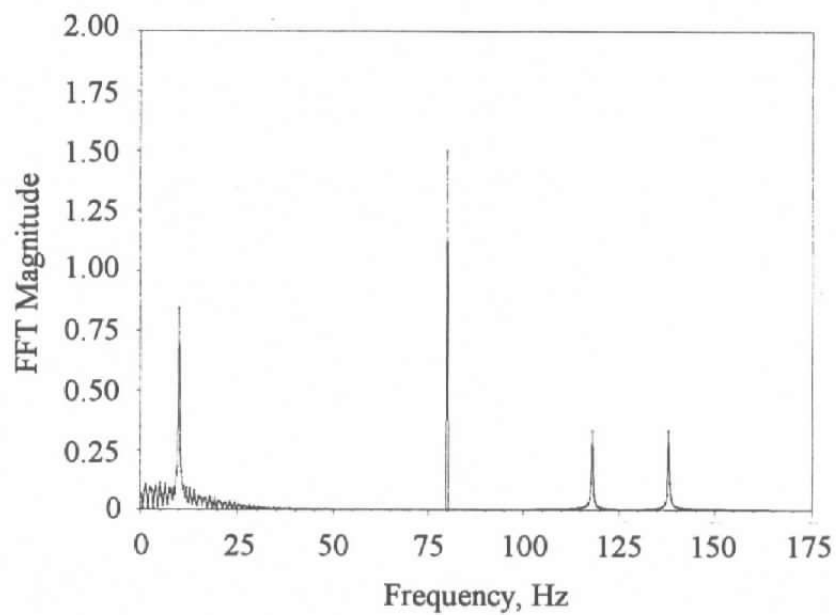
b. FFT spectrum composite signal for STFT analysis
Figure 1. STFT example signal and FFT spectrum.
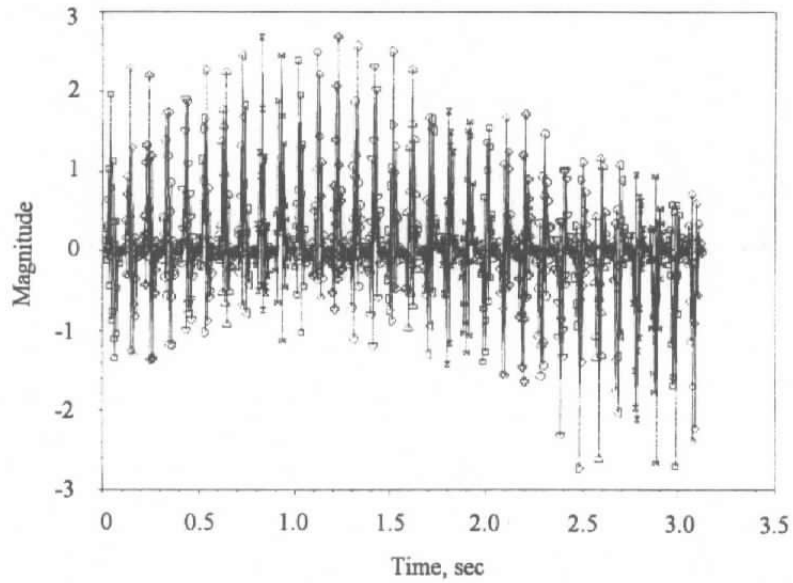
a. Windowed signal for STFT analysis (FFT Size = 32)
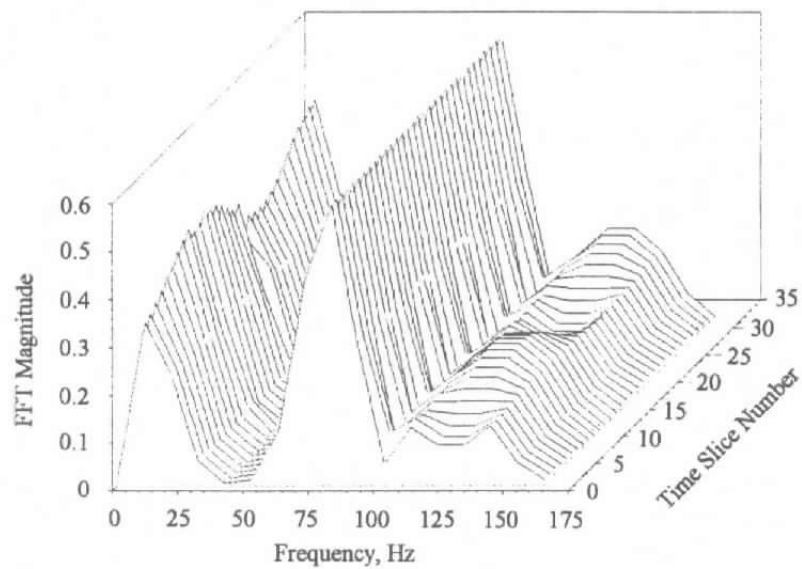
b. Composite spectrum for STFT analysis (FFT Size = 32)
Figure 2. STFT spectrogram example.
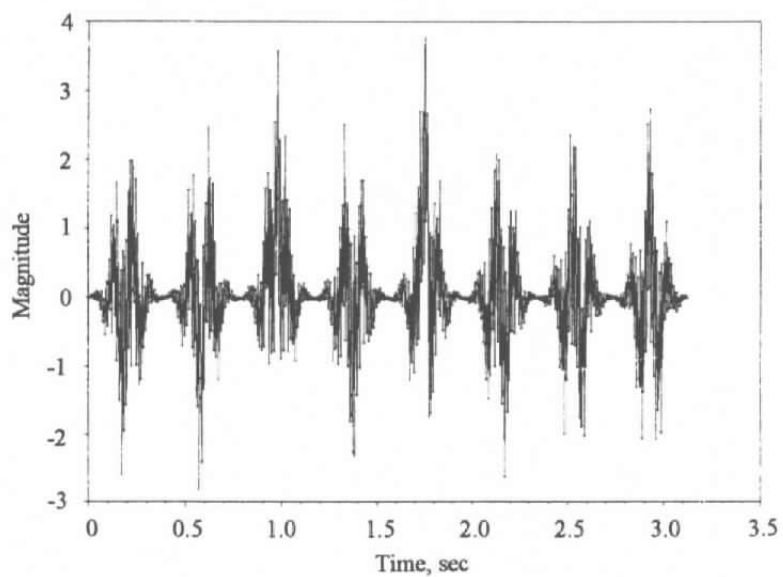
a. Composite signal for STFT analysis (FFT Size = 128)
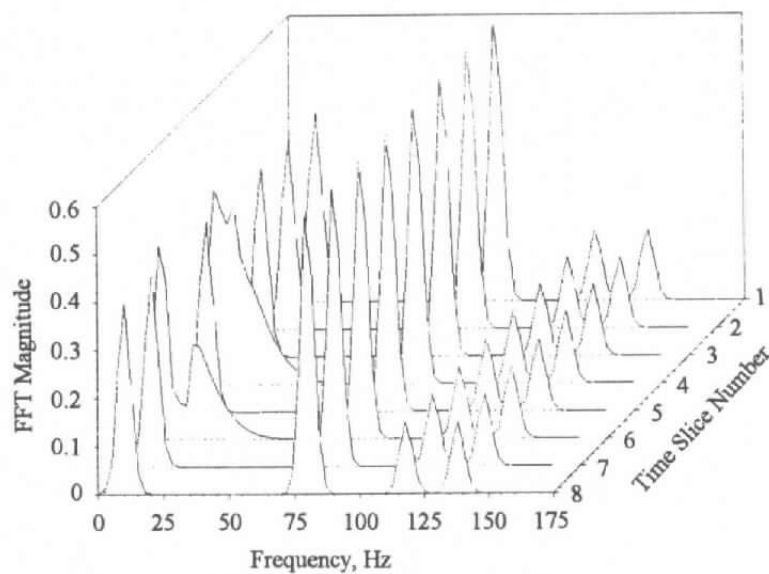
b. Composite spectrum for STFT analysis (FFT Size = 128)
Figure 3. STFT spectrogram example with improved high-frequency
time location.

41
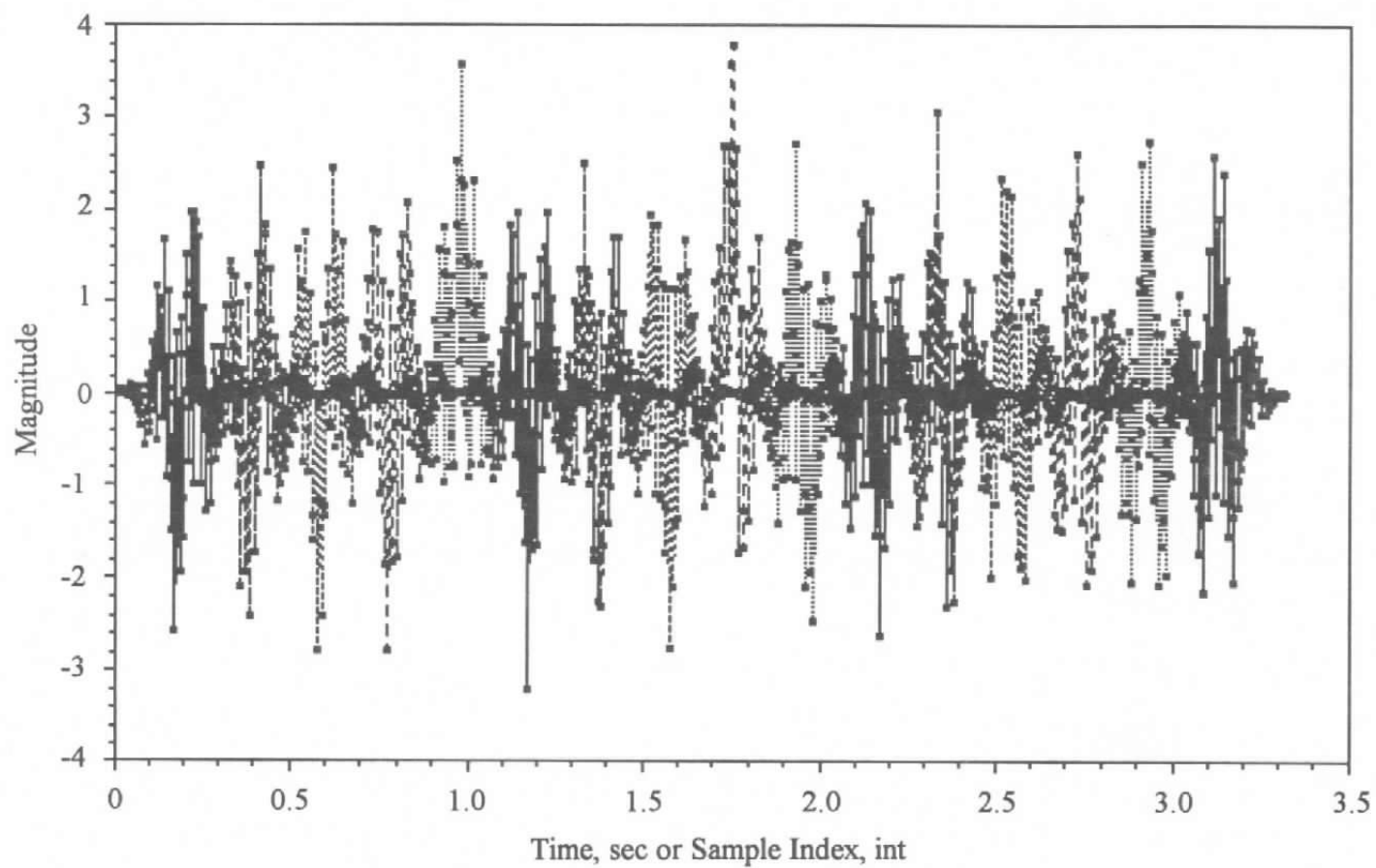
a. Composite STFT (50.00 Pct Overlap)
Figure 4. STFT windowing, 128 samples.
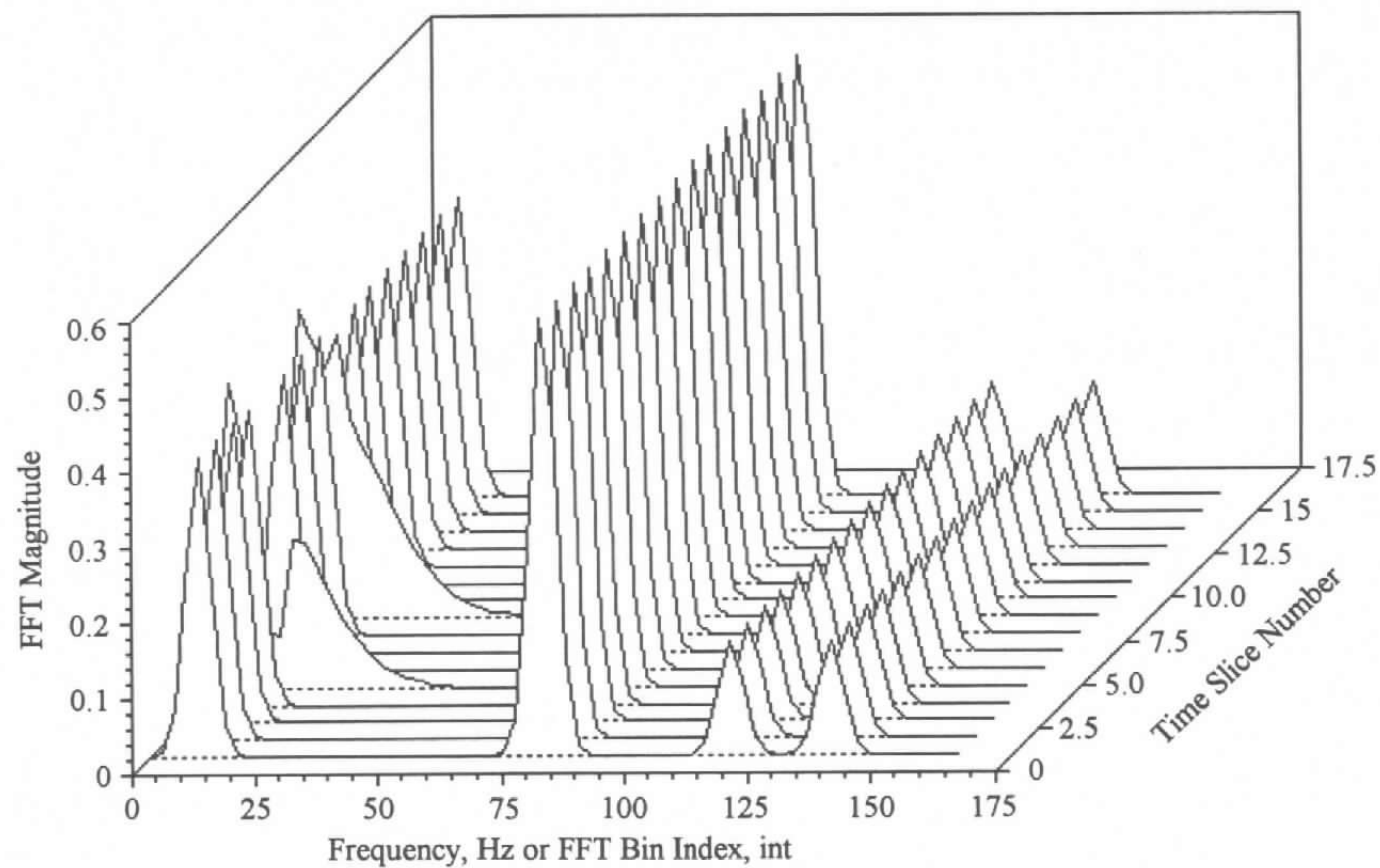
b. Composite STFT (50.00 Pct Overlap)
Figure 4. Concluded.
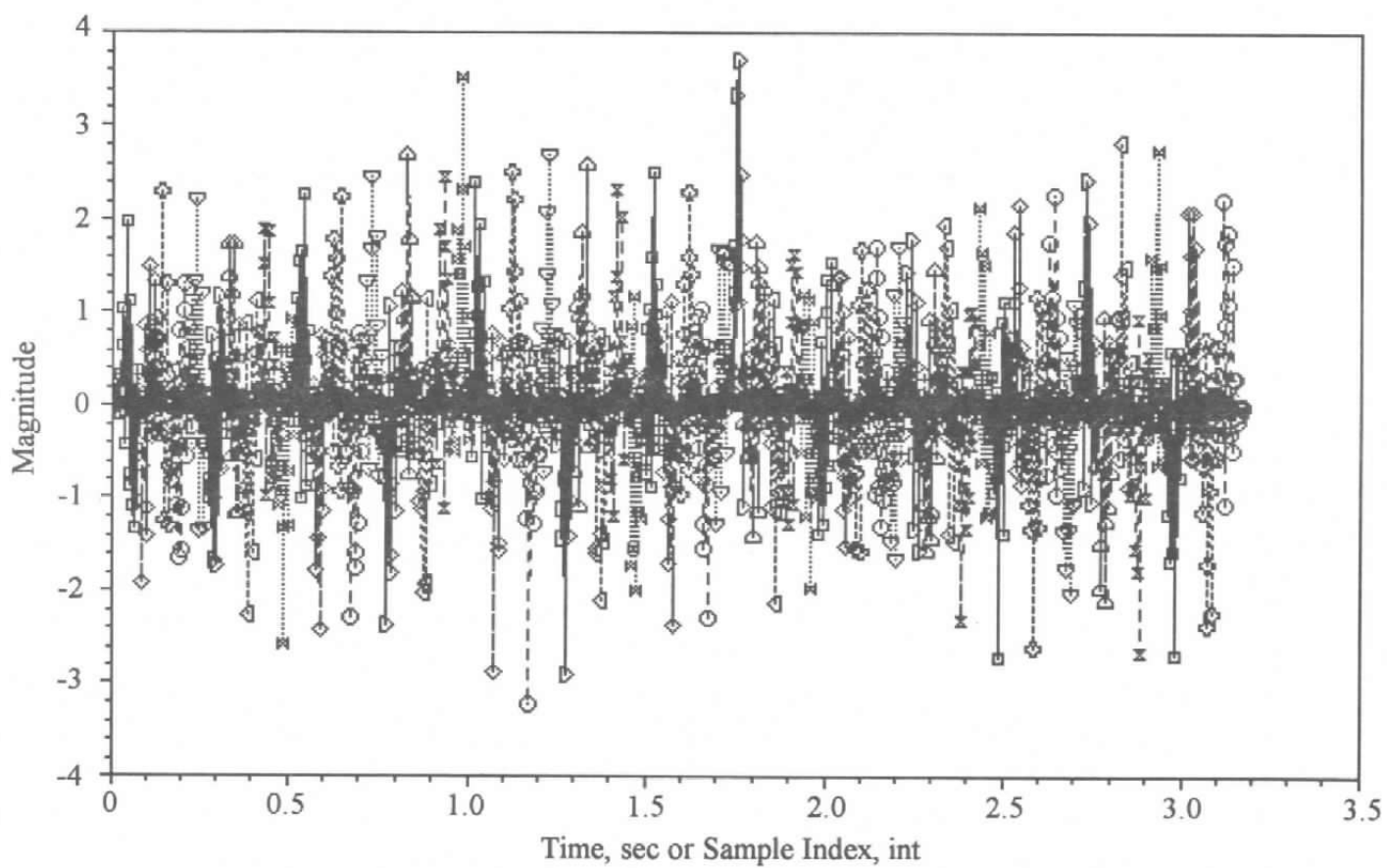
a. Composite STFT (50.00 Pct Overlap)
Figure 5. STFT windowing, 32 samples.
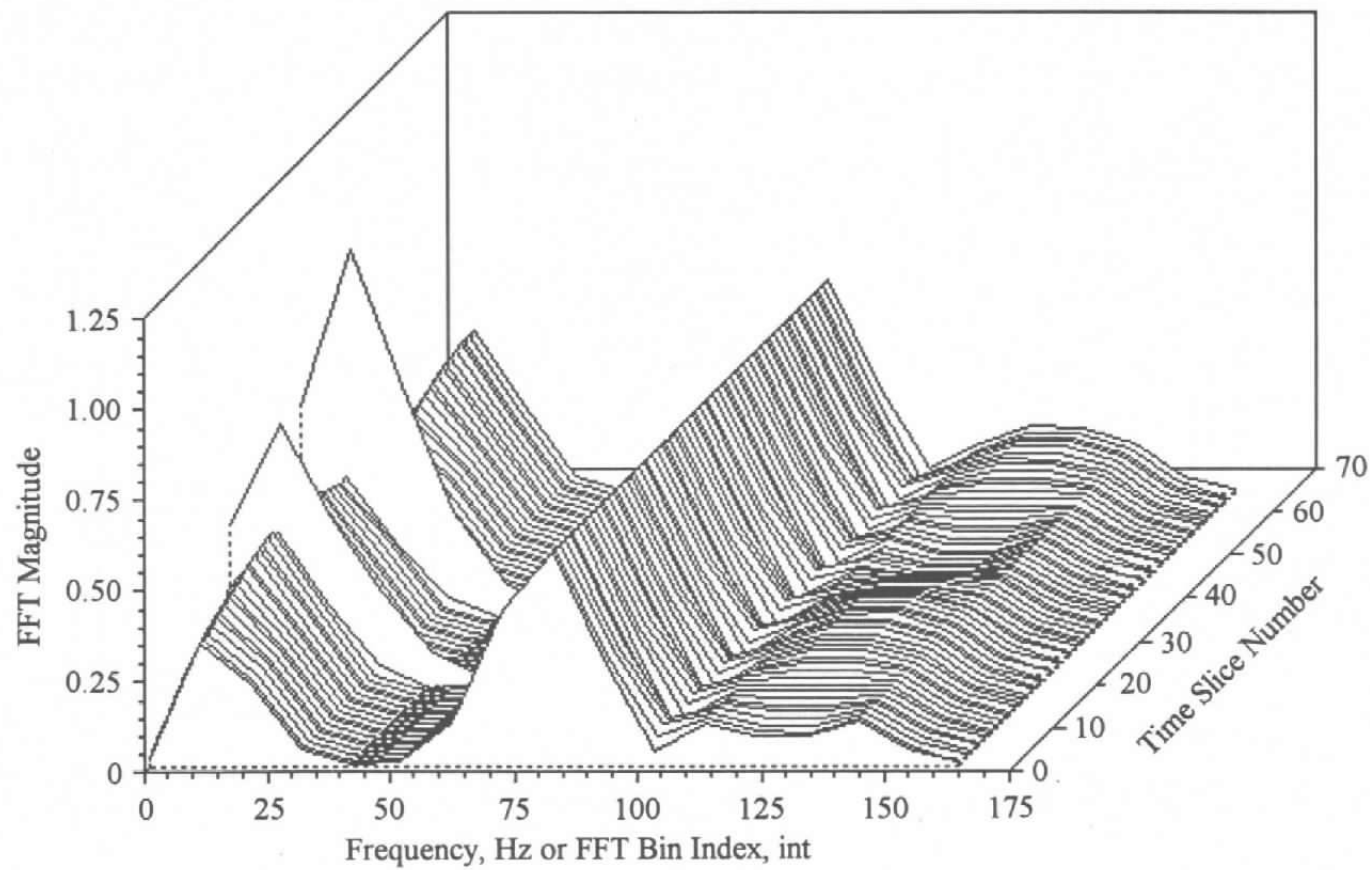
b. Composite STFT (50.00 Pct Overlap)
Figure 5. Concluded.
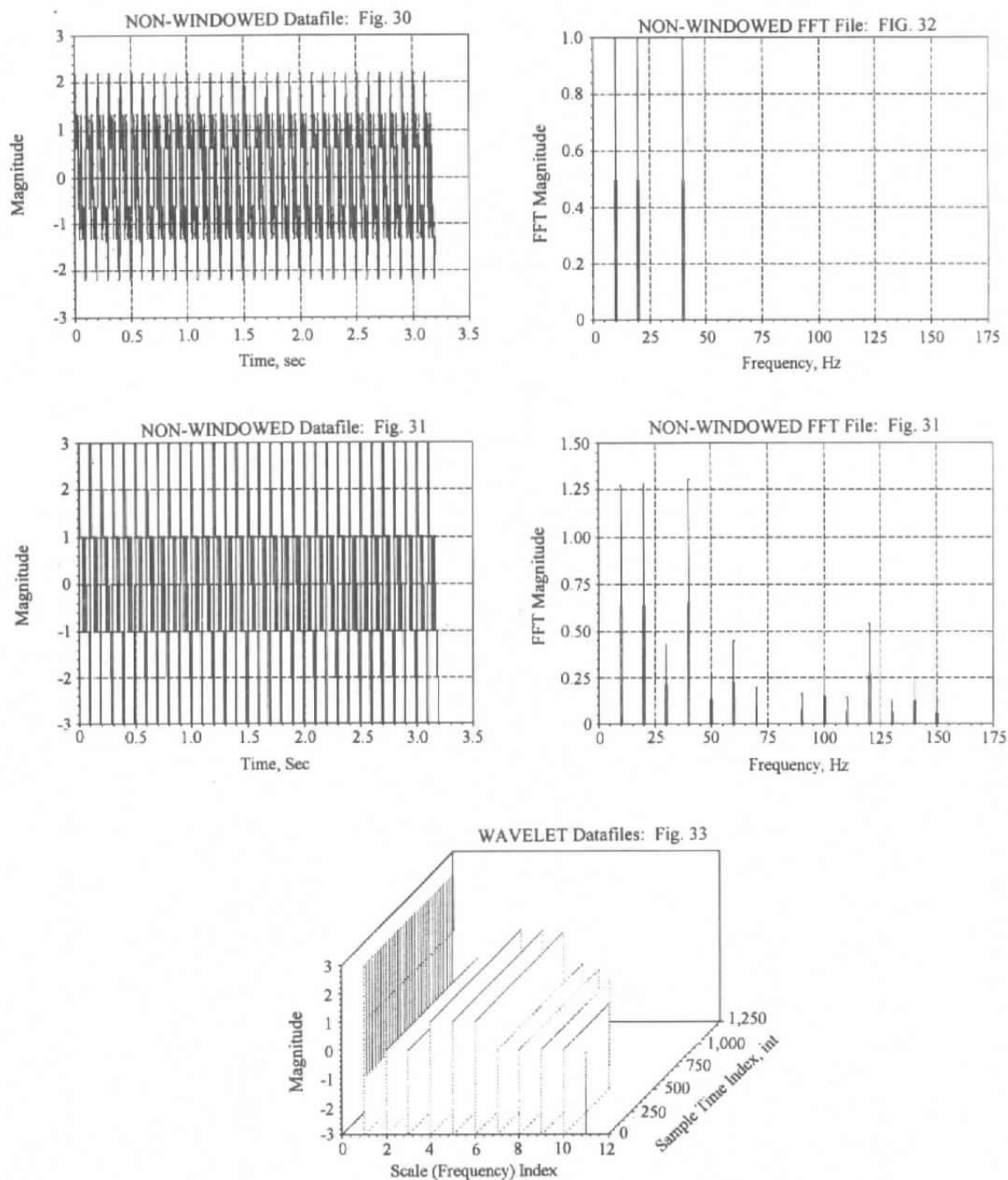
**Figure 6. Comparison of Fourier analysis for sine waves and Haar wavelet
analysis for square waves.**
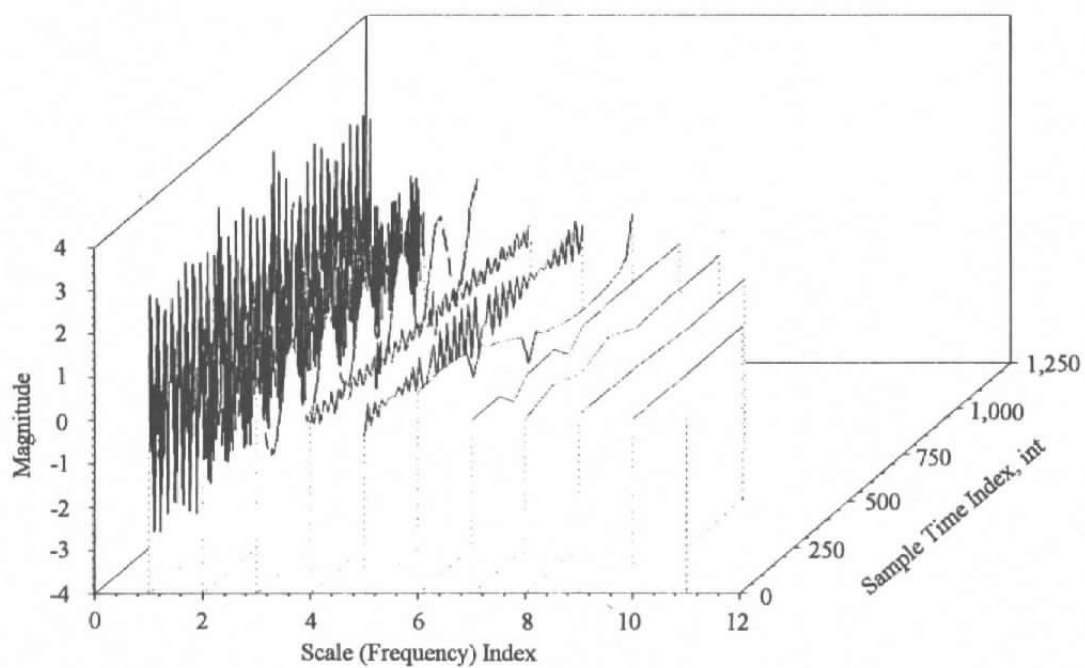
Figure 7. DAUB4 wavelet applied to STFT example signal (Ref. Fig. 14).
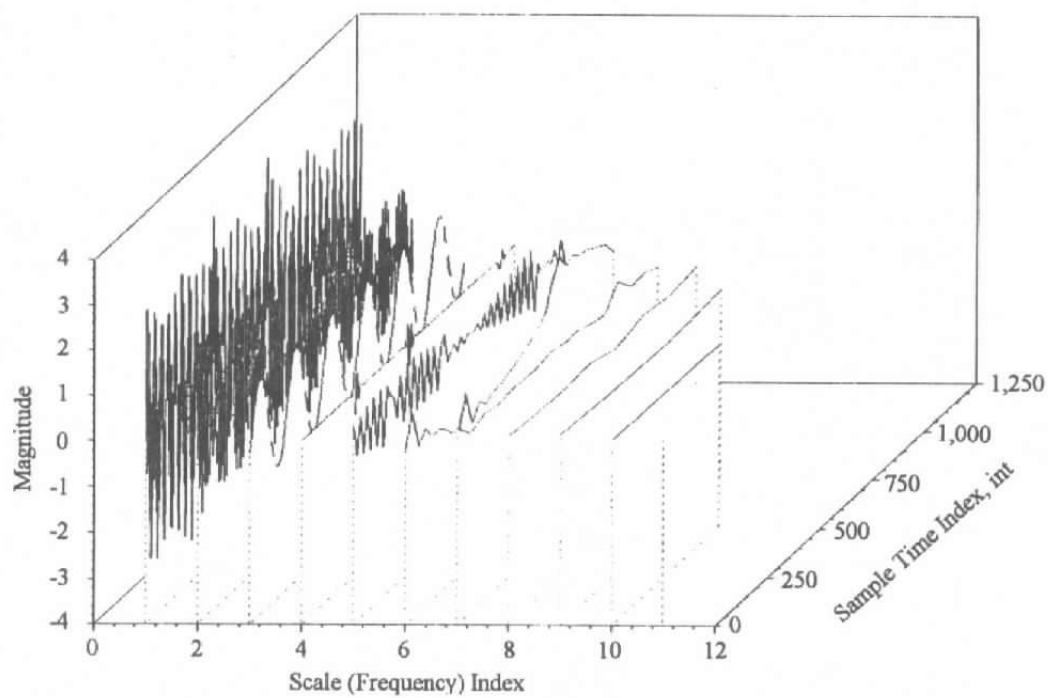
Figure 8. DAUB 20 wavelet applied to STFT example signal (Ref. Fig. 14).

48
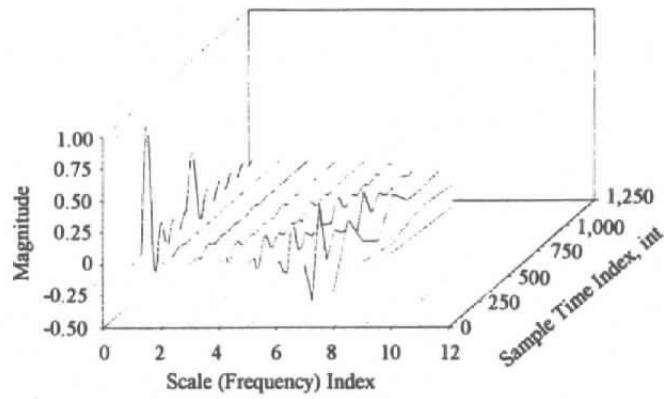
a. Wavelet datafiles, Fig. 59
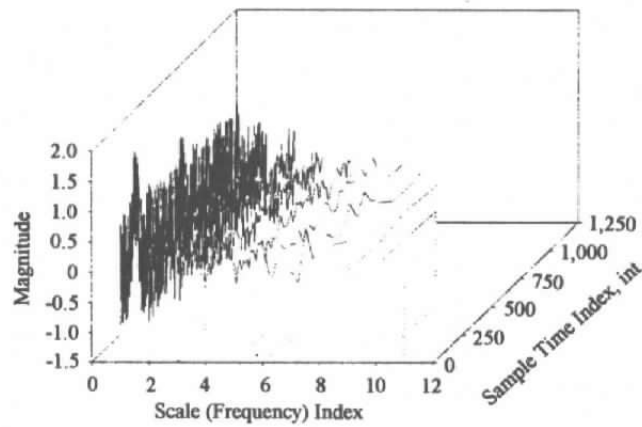
b. Wavelet datafiles, Fig. 63
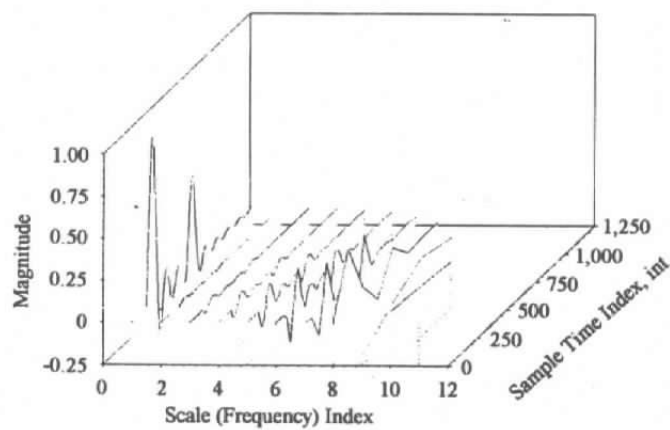
c. Wavelet datafiles, Fig. 64
Figure 9. Wavelet analysis of pulse-echo simulated data.

NON-WINDOWED DataFile: Fig. 59

NON-WINDOWED FFT File: Fig. 59
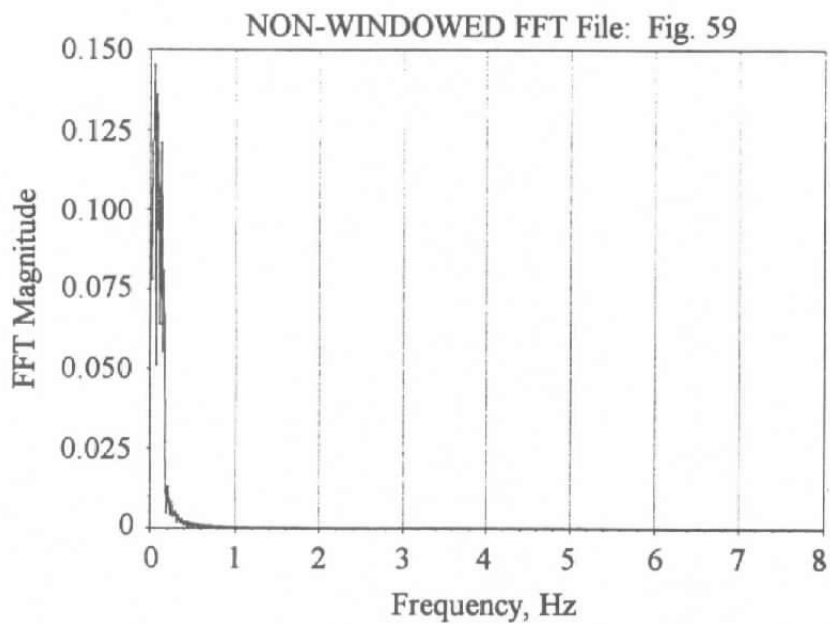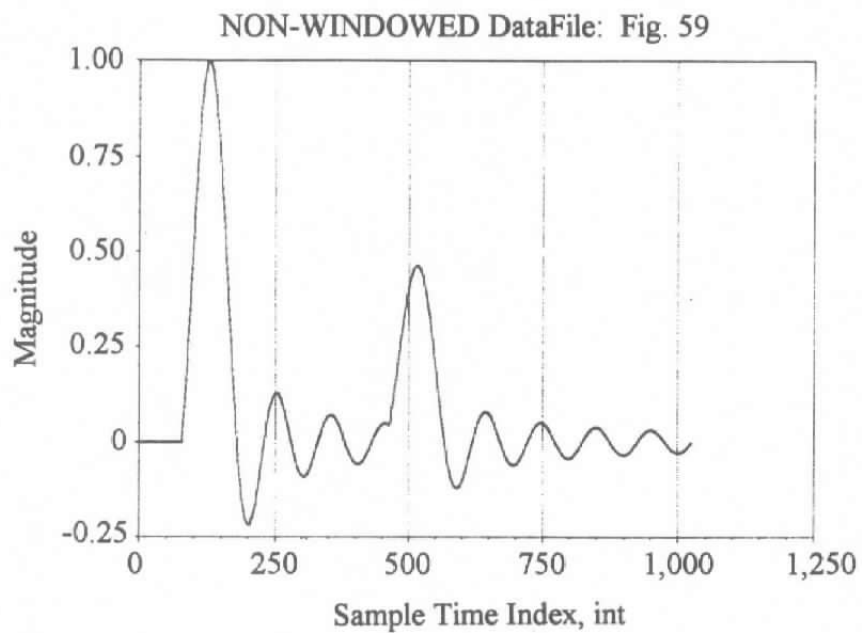
Figure 10. FFT of simulated pulse-echo data of Fig. 37a.
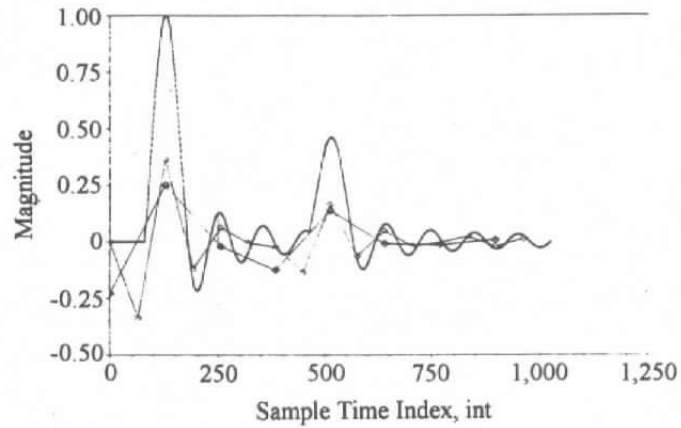
a. Wavelet detail levels 10, 4, and 3 only for Fig. 37a
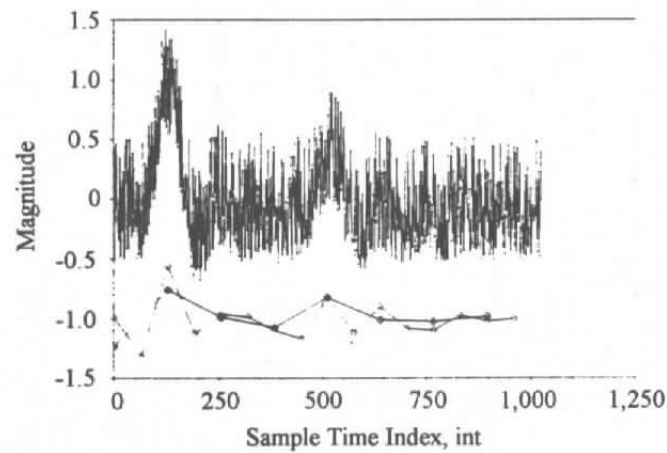
b. Wavelet detail levels 10, 4, and 3 only for Fig. 37b
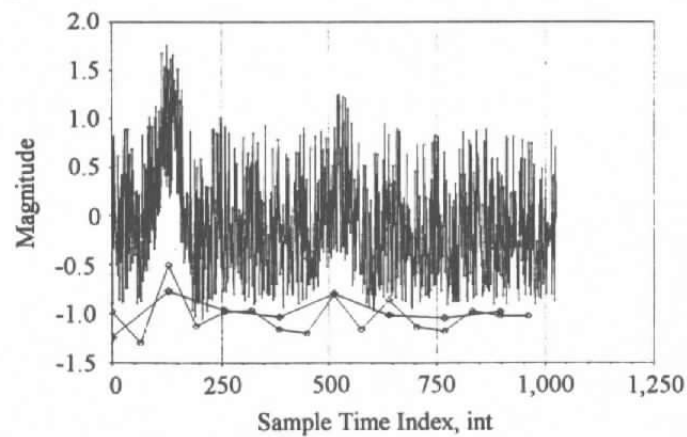
c. Wavelet detail levels 10, 4, and 3 only for Fig. 37c
Figure 11. Haar wavelet details superimposed on simulatd pulse-echo signal.

WAVELET DataFiles Fig 65

WAVELET Detail Levels 10, 4, and 3 ONLY
for Figure 40a.

WAVELET DataFiles Fig 65

WAVELET Detail Levels 10, 4, and 3 ONLY
for Figure 40b.
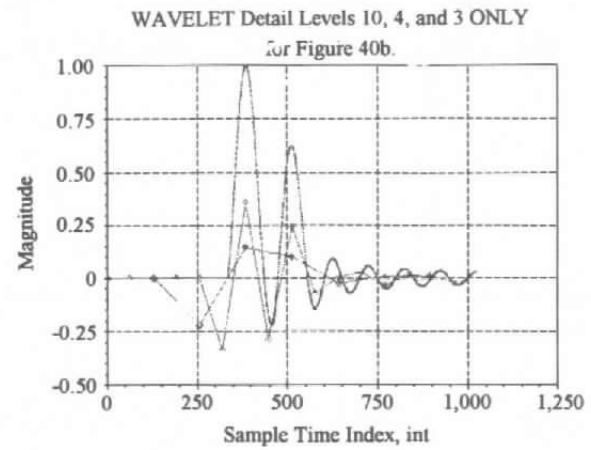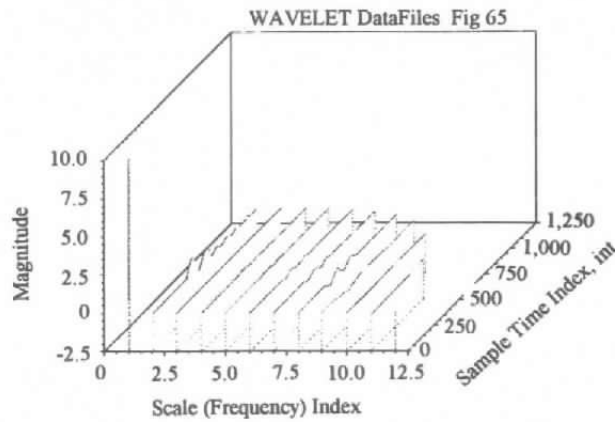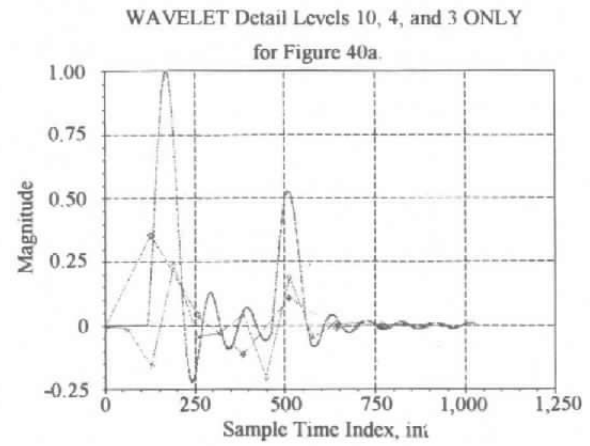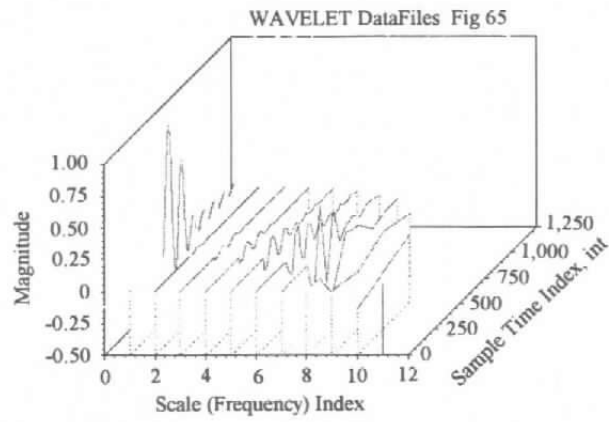
Figure 12. Wavelet analysis of pulse-echo simulated data with peak not coincident
with a 2^j sample point.
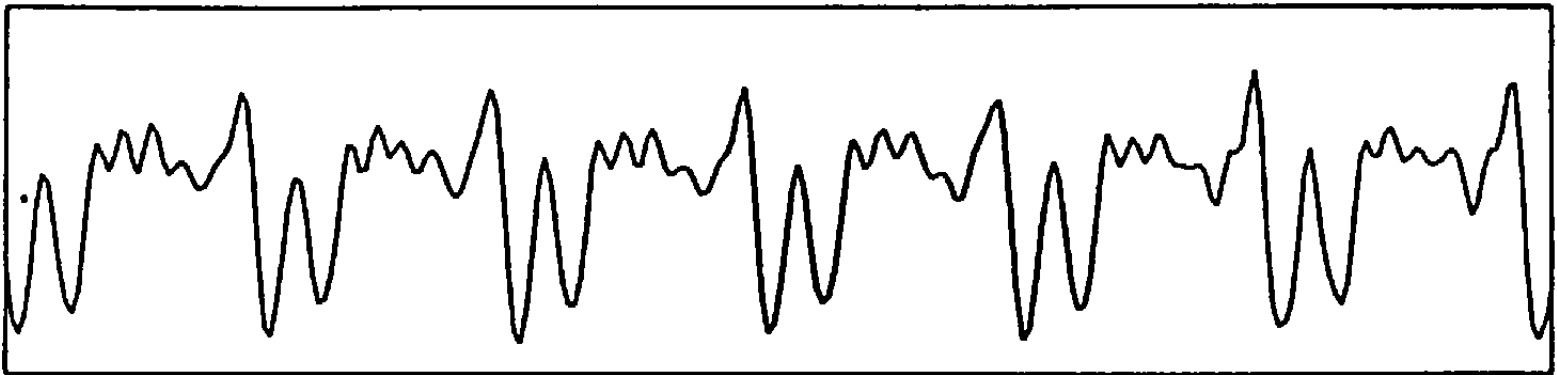
Figure 13. Acoustic signal reconstruction after discarding several least significant spectral coefficients.

Overlapping Malvar Windows
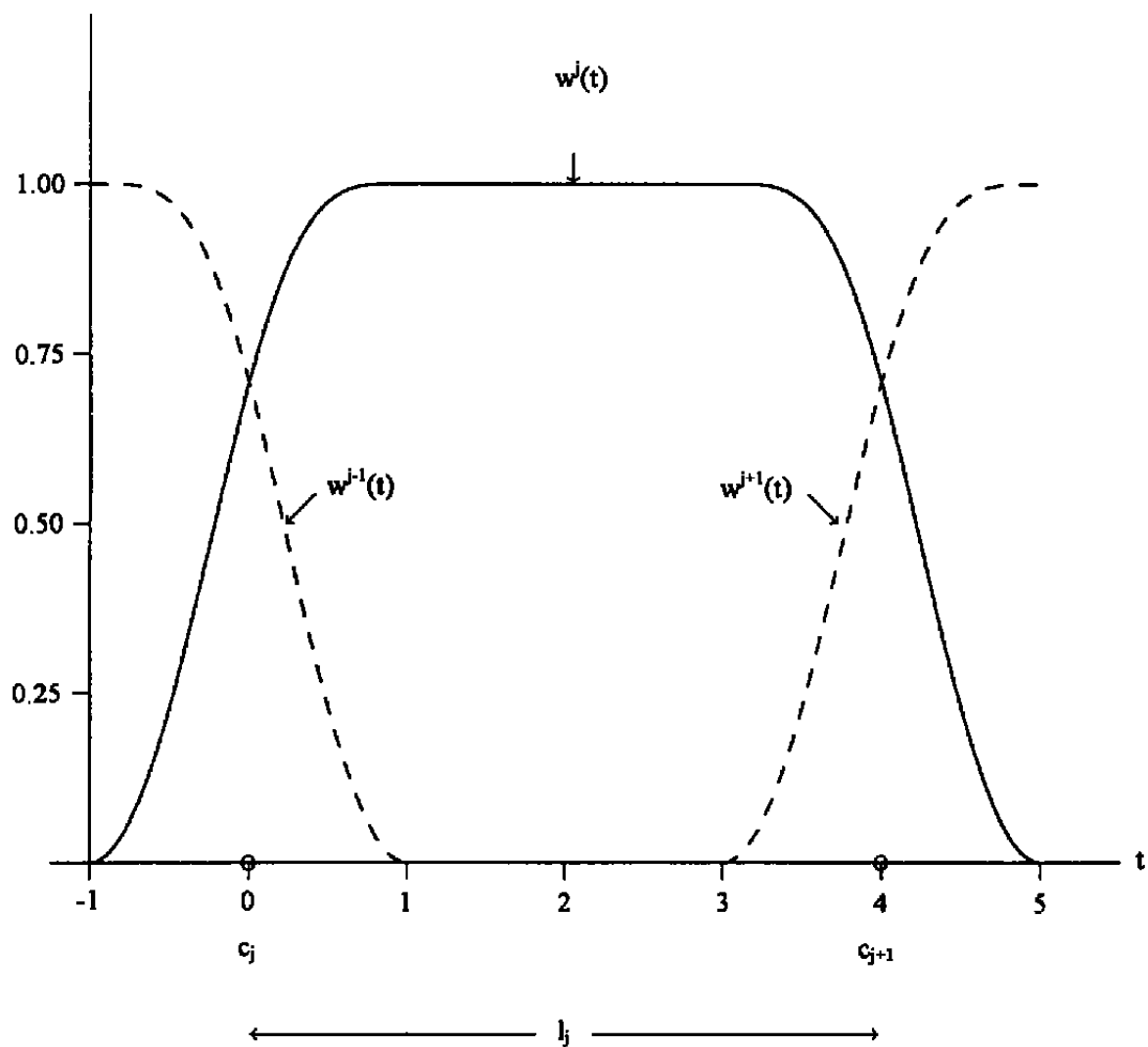
Figure 14. Overlapping Malvar windows.

$$\Psi^j_k(t) = w_j(t)\cos[(k+1/2)(t - c_j)\pi/l_j]$$
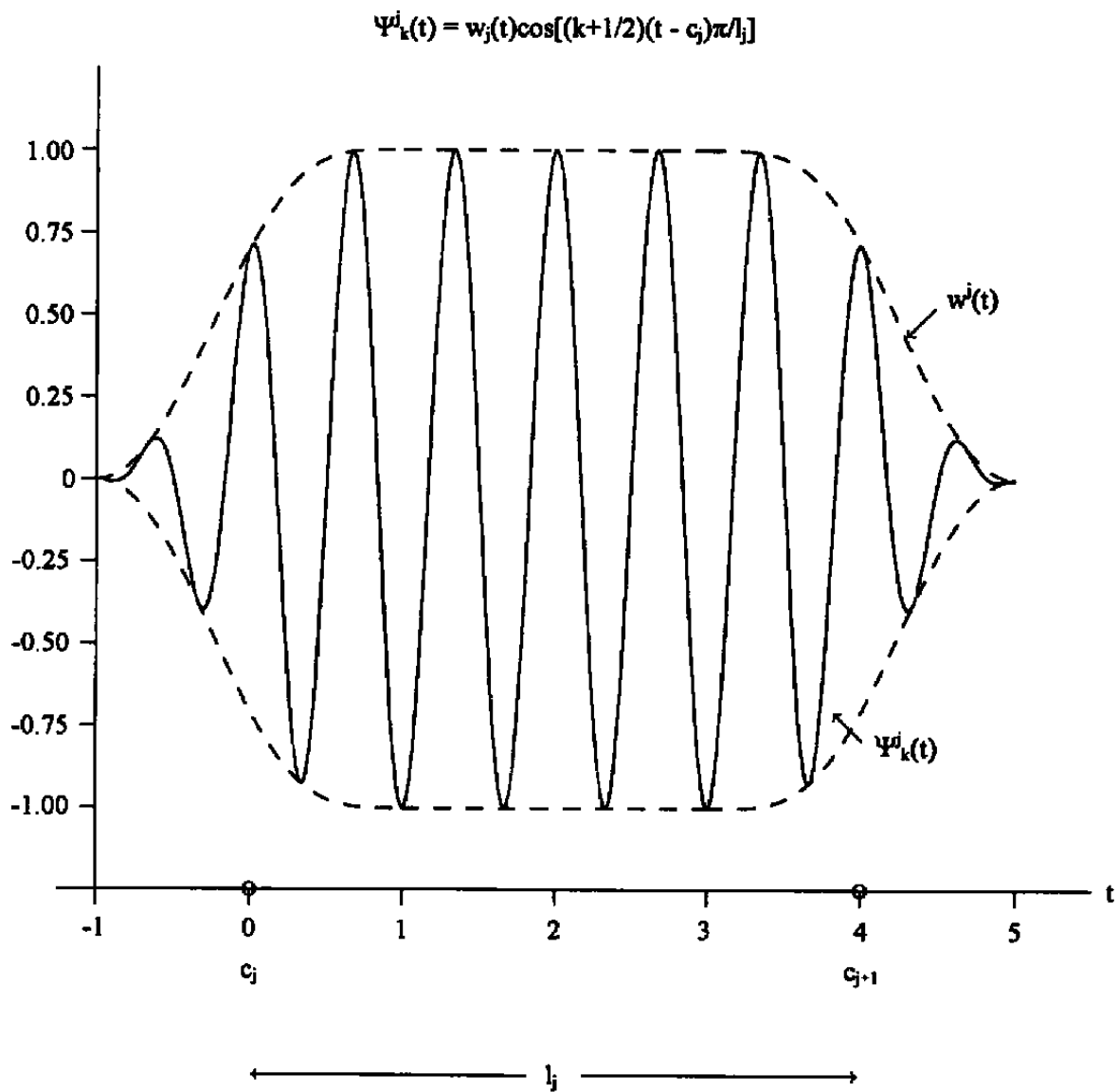
Figure 15. Typical Malvar windowed basis function.

Variable Length Windows for Malvar Transform
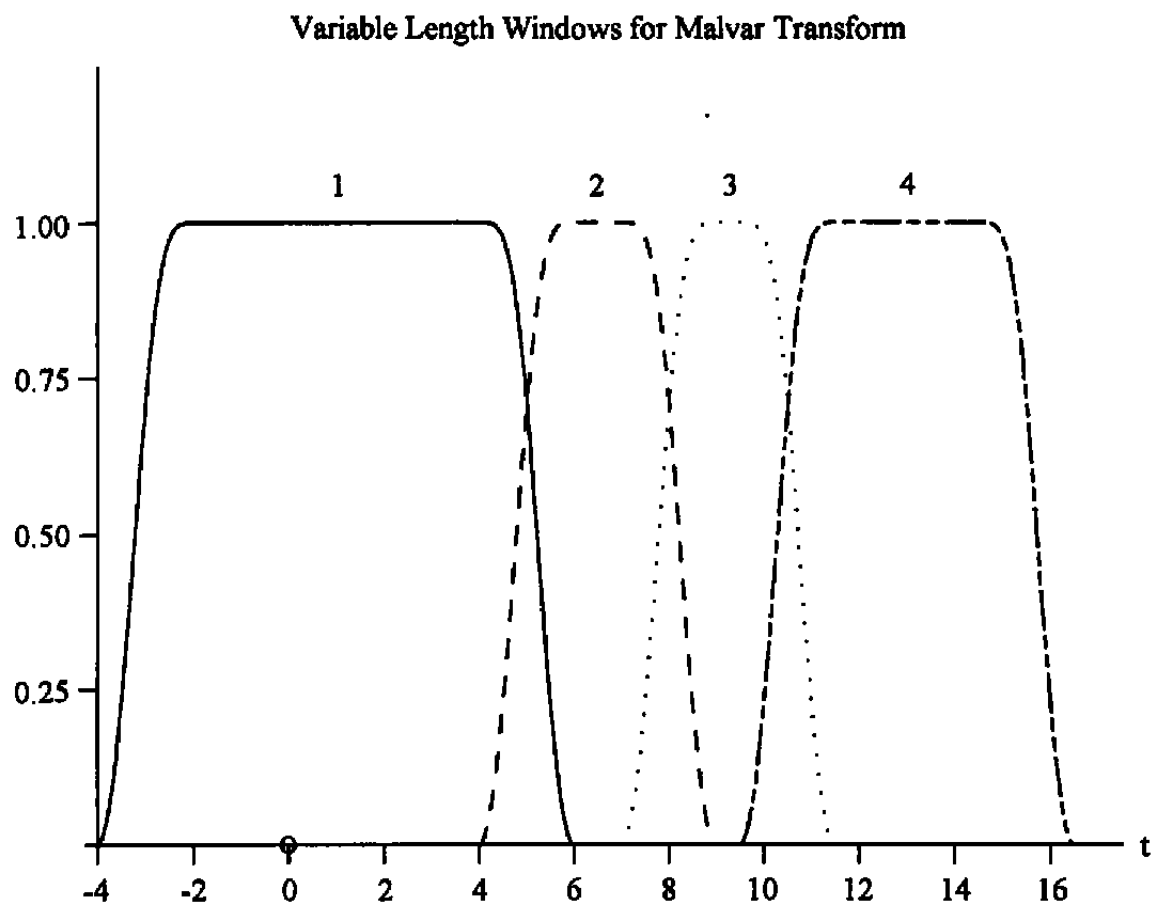
Figure 16. Variable length windows for Malvar transform.
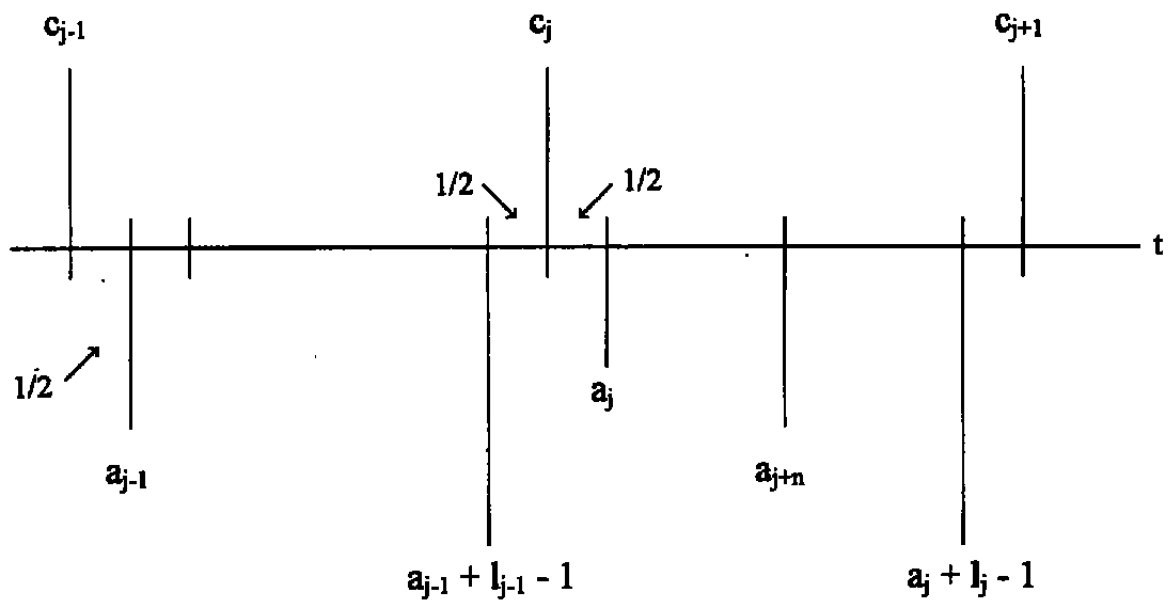
Figure 17. Discretization for Malvar wavelets.

Bell Function Profile $\beta^j(\alpha_j t) = \sin\{\pi/4[1 + \sin(\pi/2\ \alpha_j t)]\}$
where $\beta^j(t)^2 + \bar{\beta}^j(t)^2 = 1$
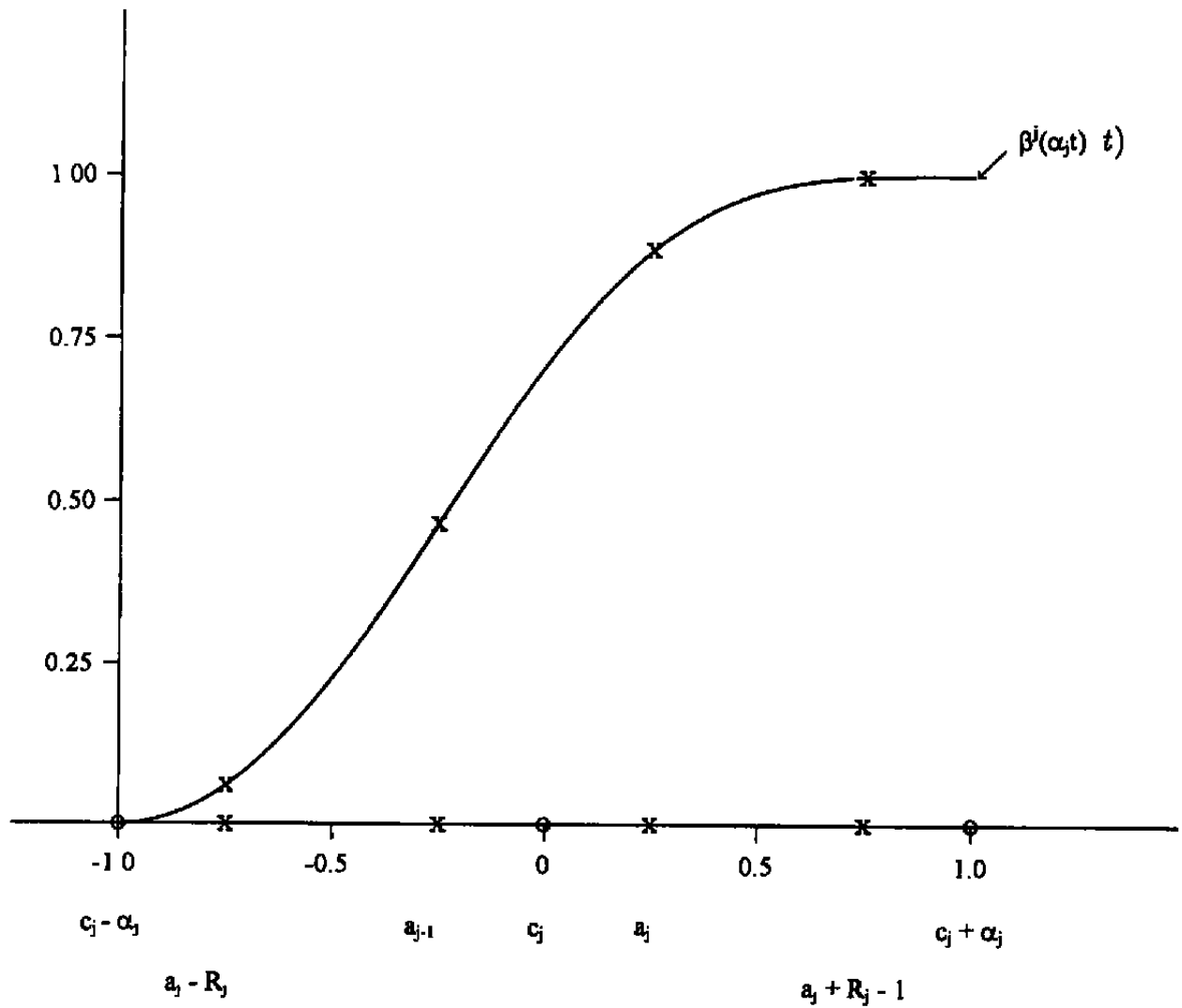
Figure 18. Bell function profile.

Comparison of Two Bases
from the Malvar Library
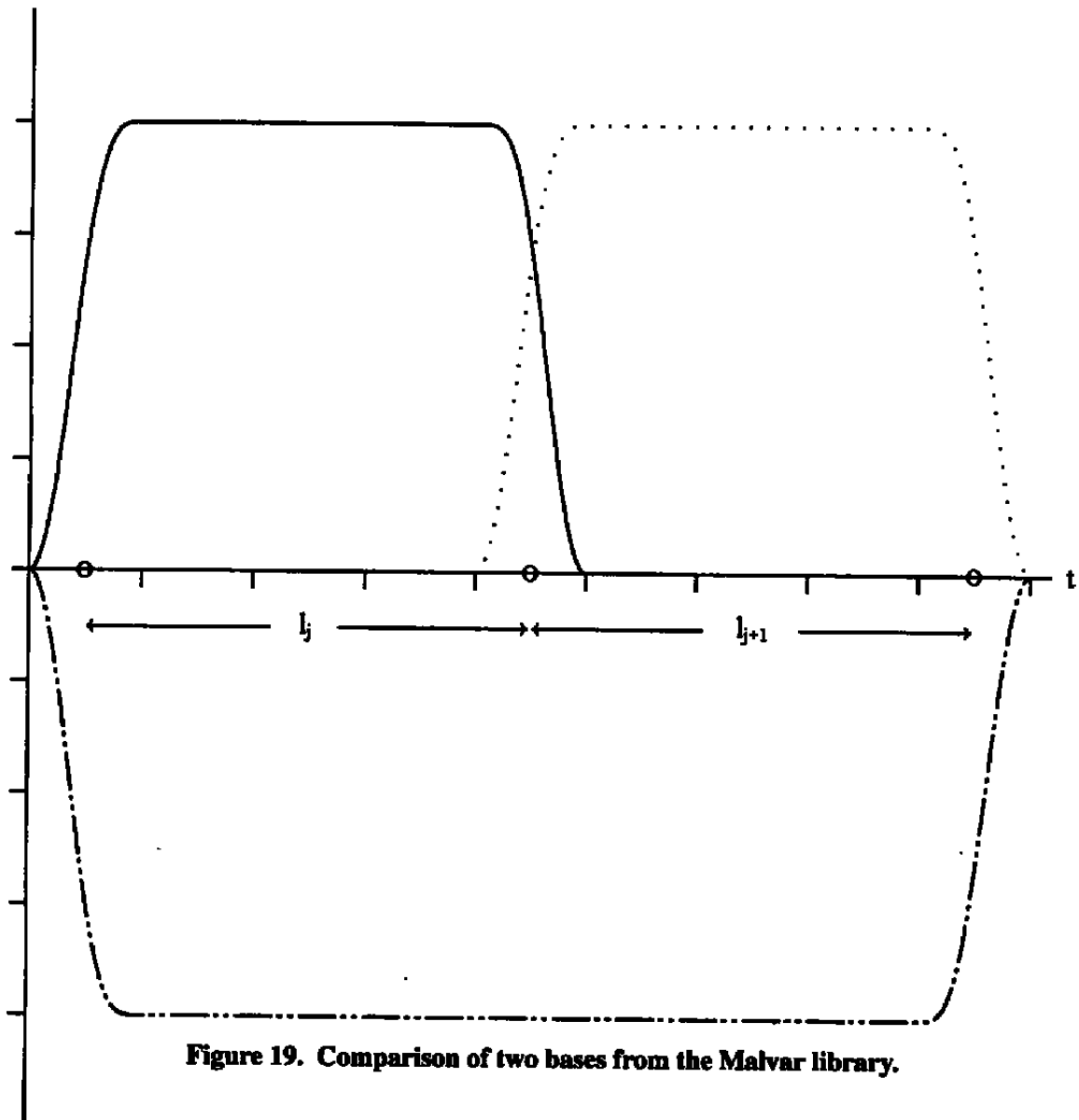
Figure 19. Comparison of two bases from the Malvar library.
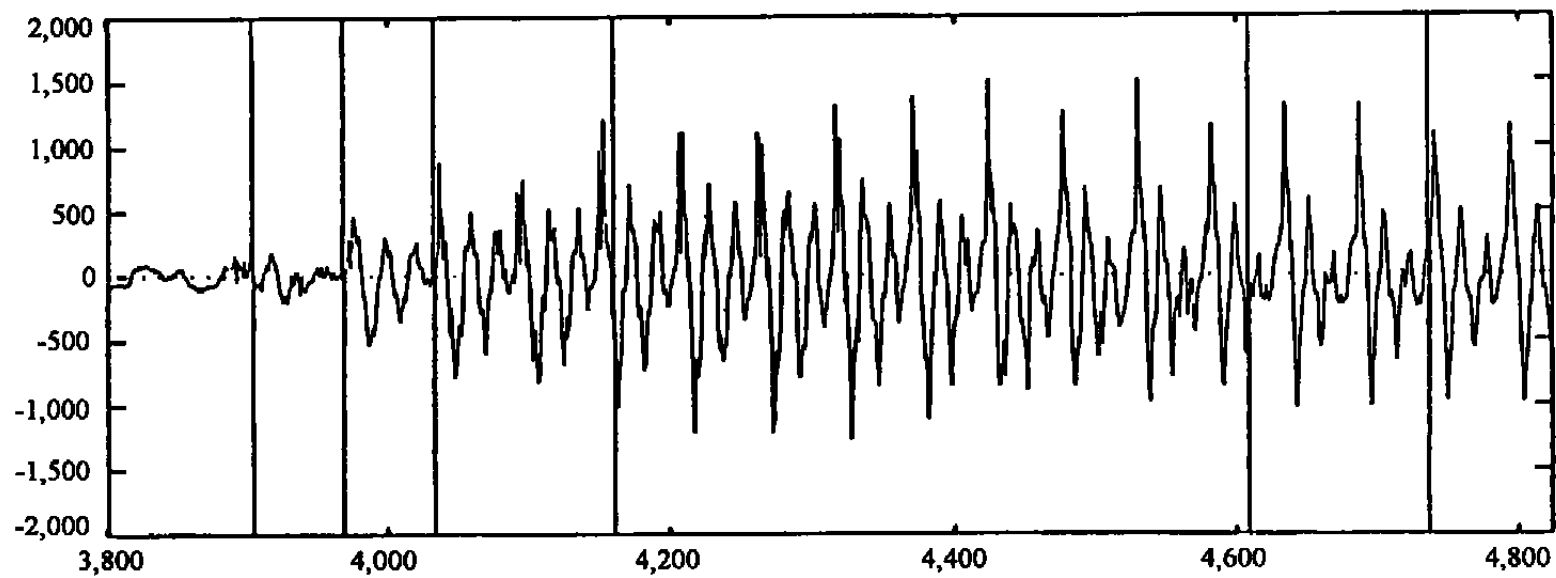
Figure 20. Automatic lowest entropy segmentation of part of a word.
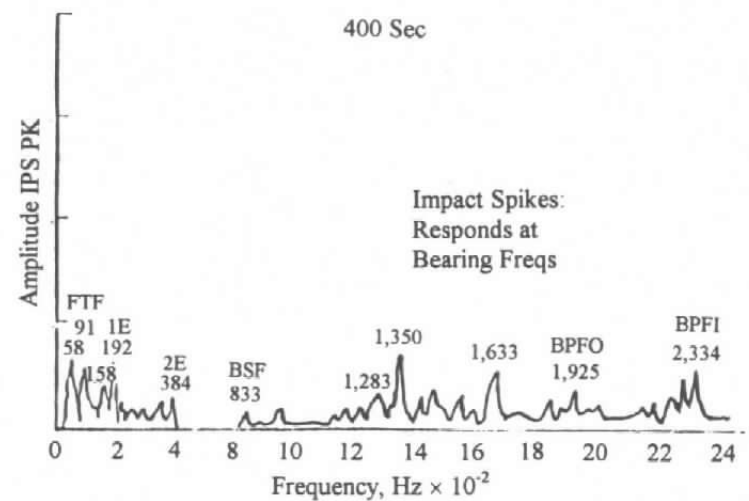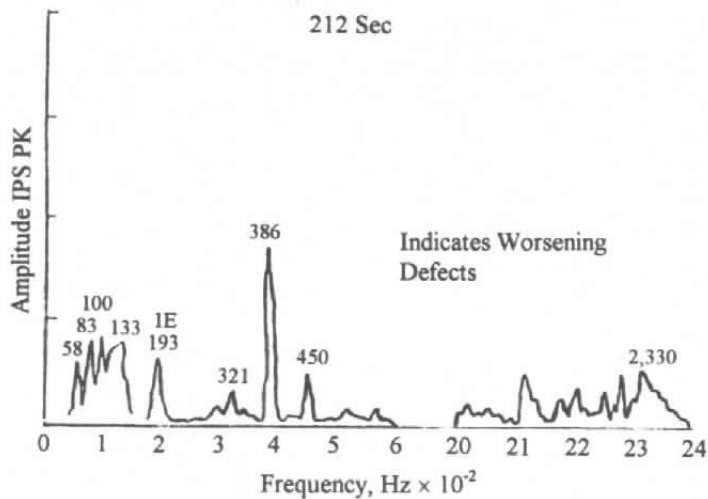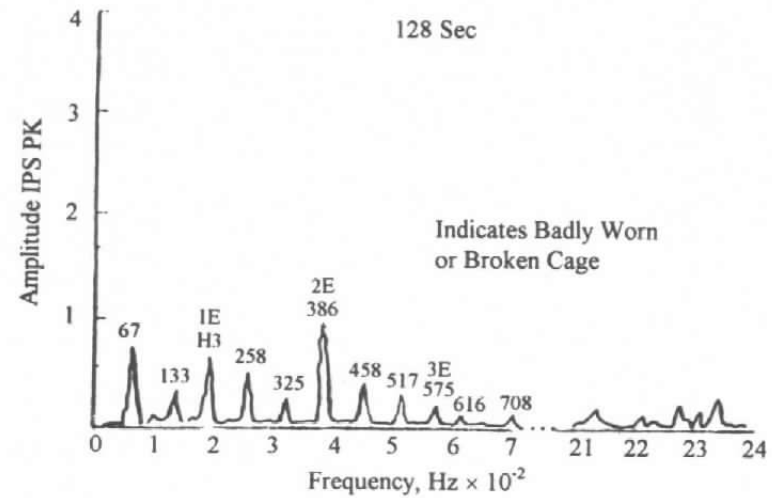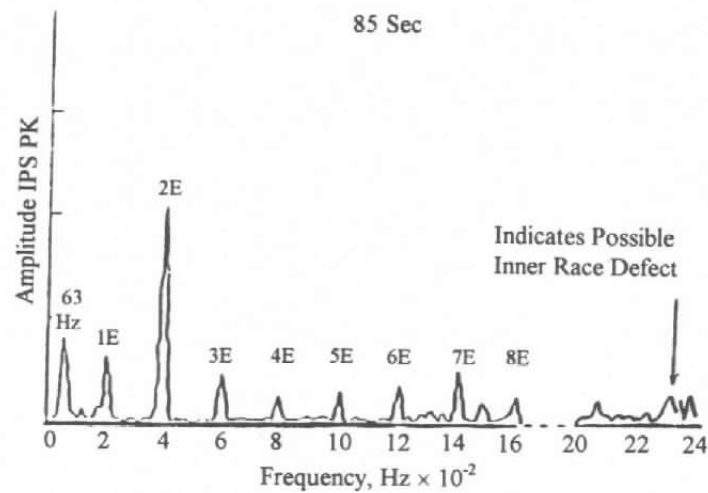
Figure 21. Typical Fourier spectra for various bearing failures.
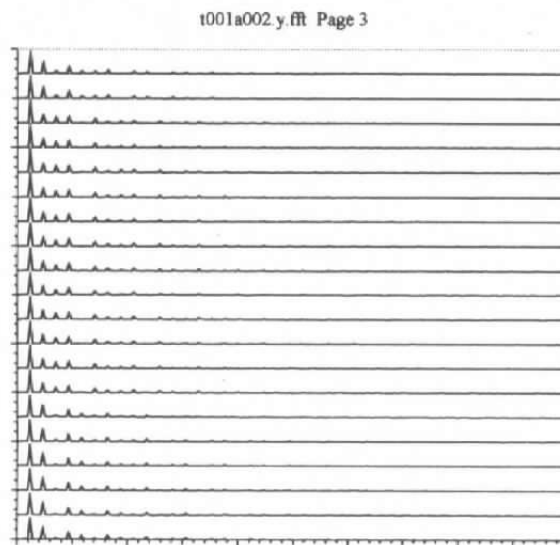
t001a002.y.sig Page 1

a. Page 1

t001a002.y.sig Page 2

b. Page 2

t001a002.y sig Page 3
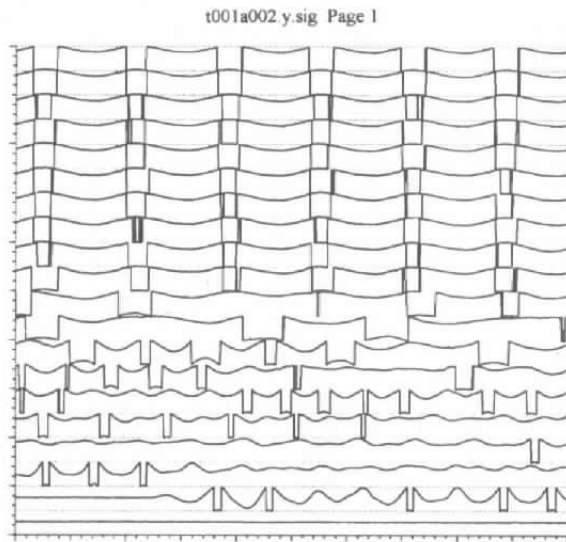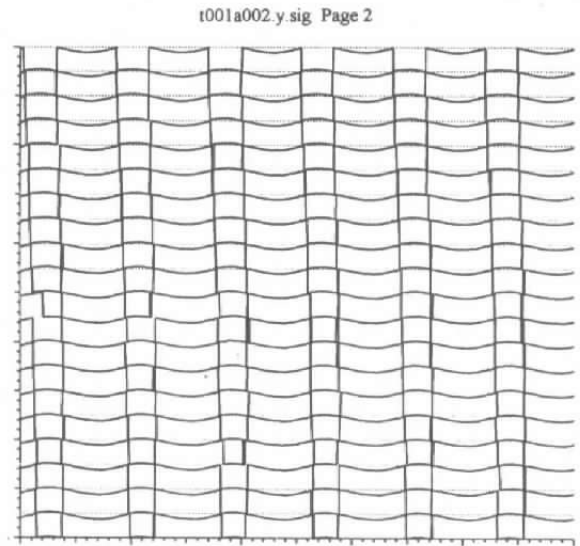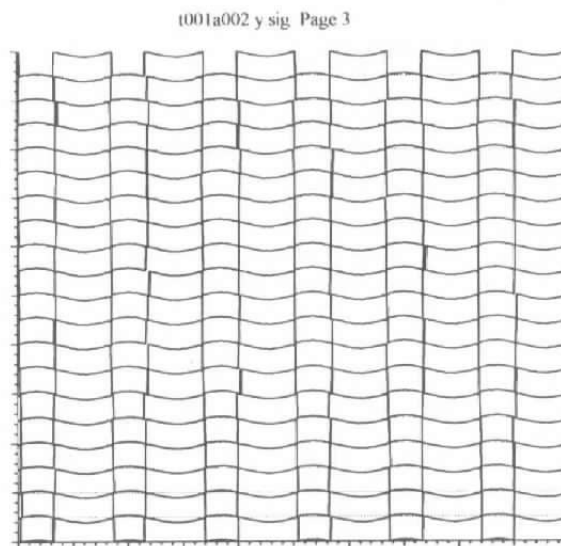
c. Page 3

Figure 22. Time domain signal for impending bearing failure.

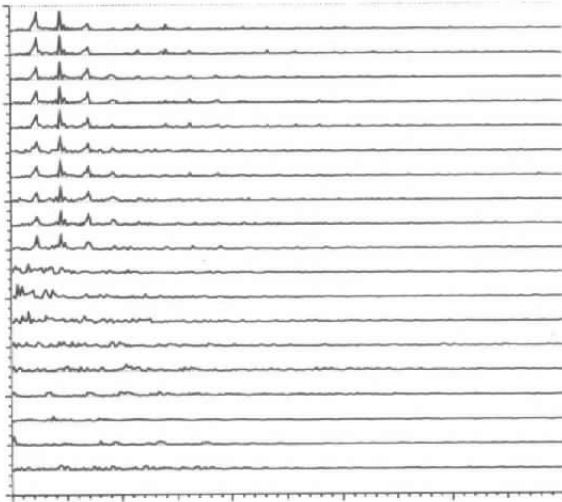t001a002.y.fft Page 1

t001a002.y.fft Page 2

a. Page 1
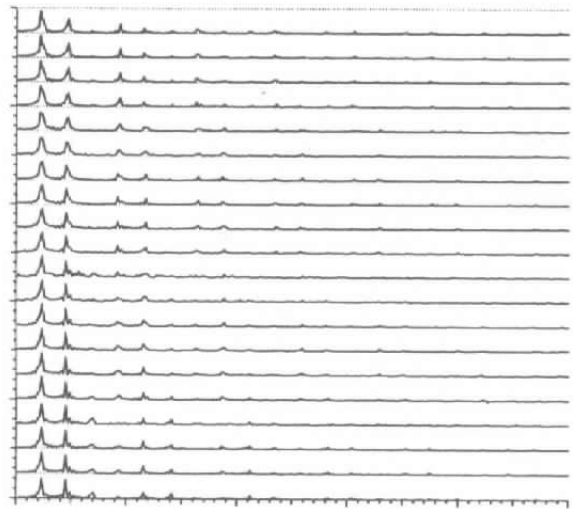
b. Page 2

t001a002.y.fft Page 3
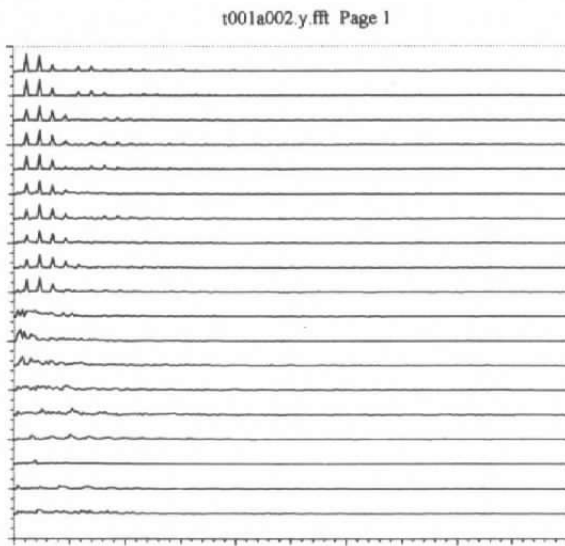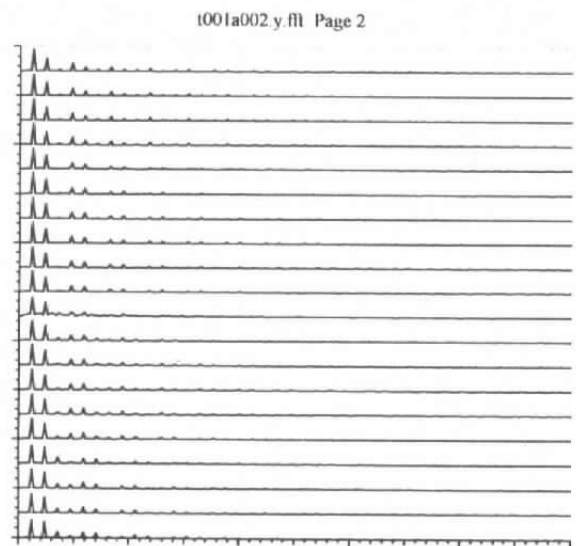
c. Page 3

Figure 23. Fourier transform for Fig. 22 signal.
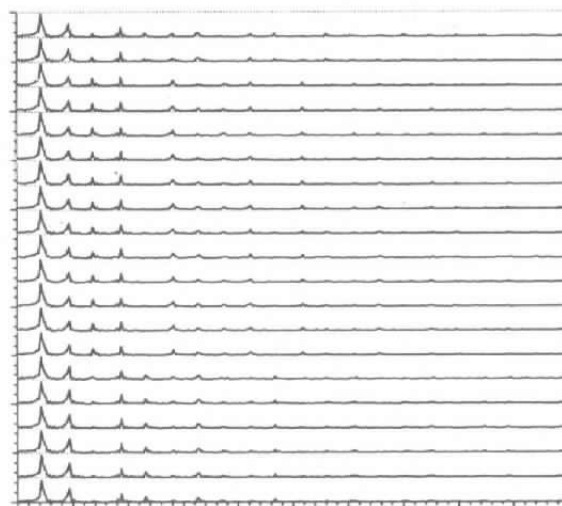
t001a002.y.mal Page 1

t001a002.y.mal Page 2

a. Page 1

b. Page 2

t001a002.y.mal Page 3

c. Page 3

Figure 24. Malvar transform for Fig. 22 signal.